# Differential Privacy Models for Location- Based Services

Ehab Elsalamouny, Sébastien Gambs

TRANSACTIONS ON DATA PRIVACY 9 (2016) 15–48

# Differential Privacy Models for Location-Based Services

**Ehab ElSalamouny**[*,**], **Sébastien Gambs**[***]

[*]INRIA, France.

[**]Faculty of Computers and Informatics, Suez Canal University, Egypt.

[***]Université du Québec à Montréal (UQAM), Canada.

E-mail: ehab.m.s@gmail.com, gambs.sebastien@uqam.ca

Received 29 May 2015; received in revised form 24 December 2015; accepted 7 February 2016

**Abstract.** In this paper, we consider the adaptation of differential privacy to the context of location-based services (LBSs), which personalize the information provided to a user based on his current position. Assuming that the LBS provider is queried with a perturbed version of the position of the user instead of his exact one, we rely on differential privacy to quantify the level of indistinguishability (*i.e.*, privacy) provided by this perturbation with respect to the user's position. In this setting, the adaptation of differential privacy can lead to various models depending on the precise form of indistinguishability required. We discuss the set of properties that hold for these models in terms of privacy, utility and also implementation issues. More precisely, we first introduce and analyze one of these models, the $(D, \epsilon)$-location privacy, which is directly inspired from the standard differential privacy model. In this context, we describe a general probabilistic model for obfuscation mechanisms for the locations whose output domain is the Euclidean space $\mathbb{E}^2$. In this model, we characterize the satisfiability conditions of $(D, \epsilon)$-location privacy for a particular mechanism and also measure its utility with respect to an arbitrary loss function. Afterwards, we present and analyze symmetric mechanisms in which all locations are perturbed in a unified manner through a noise function, focusing in particular on circular noise functions. We prove that, under certain assumptions, the circular functions are rich enough to provide the same privacy and utility levels as other more complex (*i.e.*, non-circular) noise functions, while being easier to implement. Finally, we extend our results to a generalized notion for location privacy, called $\ell$-privacy capturing both $(D, \epsilon)$-location privacy and also the notion of $\epsilon$-geo-indistinguishability recently introduced by Andrès, Bordenabe, Chatzikokolakis and Palamidessi.

**Keywords.** Location privacy, Differential privacy, Location-based services, Symmetric mechanisms, Noise functions, Geo-indistinguishability.

## 1 Introduction

The advent of ubiquitous devices equipped with positioning capacities such as smartphones has led to the growing development of location-based services (LBSs). A classical type of LBS is one providing information to a user relative to his current position such as the closest restaurants, subway stations, hospitals or even public toilets. Typically, a user accesses an LBS through his personal device. More precisely, his device submits his request

to the LBS provider together with the geographical position of the user, which is acquired by the device itself (*e.g.*, through the GPS system). Afterwards, the LBS provider personalizes the service based on the location of the user.

However, for a user revealing his location to the LBS provider may impact his privacy. In particular, the mobility data of an individual can be used to learn the points of interests characterizing his mobility such as his home and place of work, to predict his past, current and future locations or even to discover his social network. This privacy issue has initiated a trend of research whose objective is to allow individuals to use such services while protecting their location privacy. An early and yet intuitive approach is to remove the identity of a user from his request and to replace it with a pseudonym [1]. However, it turned out that it is often possible for an adversary having some background knowledge about users (possibly obtained through a public source) to de-anonymize a set of pseudonymized mobility traces [2]. Thus in addition of hiding his identity, it is also important to obfuscate the position of a user. For instance in [3, 4], the location of a user in a request is $k$-anonymized before being released to the LBS provider, which results in his mobility behavior being indistinguishable from $k-1$ other users. In this context, the privacy level provided depends on the background knowledge of the adversary, which could be in fact the LBS provider himself [5].

In this paper, we are interested in the models of location privacy that abstract away from the background knowledge of the adversary. In particular, we would like to have privacy guarantees holding for the user based only on the properties of the mechanism obfuscating his location. One intuitive possibility is to divide the space into zones (or cells) and query the LBS by sending him the identifier of the zone that encloses the user's real location.

While this solution is appealing in some situations (*e.g.*, in the case of repeated queries inside the same zone), it endangers the location privacy in other cases. For example the home and working place of a user may be arbitrarily close to each other while being in different zones. In this scenario, reporting the enclosing zone when the user is at one of these two places enables a curious adversary to exclude the possibility of the other location (*i.e.*, the adversary can distinguish with certainty between these two points of interest). This violates the user's location privacy provided that he requires these two places to be indistinguishable to a certain extent.

To avoid this issue, we require the privacy guarantee to restrict the distinguishability not only between the points inside the same zone, but also between the points of different zones. In other words, we require this guarantee to hold for *all the points* of the space considered, in addition of being independent of the adversary's prior knowledge. This objective is similar in spirit to the work initiated by Dwork on *differential privacy* [6] in the context of statistical databases. In a nutshell, the main idea of differential privacy is that the presence (or absence) of an individual in the database should have a negligible impact on the probability of each output of a computation (*e.g.*, a statistical query). More precisely, (the log of) the ratio between the probabilities of obtaining a certain answer, from any two *adjacent* databases (*i.e.*, differing only in the presence of an individual), has to be lower than a given parameter $\epsilon$. As a consequence, the distinguishability between any two adjacent databases is bounded up to a certain level quantified by $\epsilon$.

Differential privacy can be extended to the context of LBSs by considering the user location as the sensitive information to be protected. More precisely, we assume that the personal device of the user applies an *obfuscation mechanism* that takes the user location as input and produces a perturbed version of the position $\boldsymbol{p}$ as output. Afterwards, the user queries the LBS provider with $\boldsymbol{p}$ instead of his real location. In this setting, the location privacy of user is protected - in the sense of differential privacy - by restricting the impact

on the probability distribution over the outputs of the mechanism when changing its input from one location to another nearby position, making these locations distinguishable from each other only to a certain extent. Depending on the exact notion of distinguishability chosen between two arbitrary positions, several models of location privacy can arise.

For instance, in this paper we introduce the notion of ($D$, $\epsilon$)-location privacy that results from adapting the adjacency relation in the standard differential privacy to the domain of locations. Precisely, two locations are considered "adjacent" if the distance between them is less than a predefined value $D$. In this context, we say that a mechanism satisfies ($D$, $\epsilon$)-location privacy if the (log of) the ratio between the probabilities of obtaining a certain output, from any two adjacent locations is at most $\epsilon$. This property guarantees that the distinguishability between the location of the user and all the points that are adjacent are always restricted to a certain level quantified by $\epsilon$.

Another possible model resulting from the adaption of differential privacy to LBSs is the concept of $\epsilon$-geo-indistinguishability [7]. In this model, the bound on the distinguishability between two arbitrary positions increases linearly with the distance $d$ between them. This means that the (log of) the ratio between the probabilities of obtaining a certain output from two locations is at most $\epsilon d$, which provides a low level of distinguishability (*i.e.*, high privacy) between neighboring positions. In contrast, a higher level of distinguishability (*i.e.*, low privacy) occurs for points that are further apart.

These two notions are not exhaustive and other location privacy models can emerge by considering other type of distinguishability notions. Abstracting from the different possible models, our main objective is to provide and analyze the main properties that should hold for location privacy models based on differential privacy in terms of privacy, utility and implementation issues. In this analysis, we assume that a mechanism takes as its input the location of the user and produces as output another location in the Euclidean planar space $\mathbb{E}^2$ using a randomized process. In this setting, we formulate location privacy as a set of constraints on the conditional probabilities of the outputs of the mechanism given arbitrary inputs. While the input domain of a mechanism is often defined as an arbitrary set $\mathcal{X}$ of locations, we assume that its output domain is the entire $\mathbb{E}^2$ plane. Indeed, typically the output of the mechanism can be remapped to another domain (*e.g.*, latitude and longitude coordinates or physical addresses) using some post processing step, which can be performed either directly on the device of the user or on the server side by the LBS provider (*e.g.*, if this remapping depends on the service).

In addition to the privacy guarantees, our analysis also addresses the *utility* (or equivalently the *expected loss*) incurred by the use of the obfuscation mechanism. More precisely, our analysis relies on an "arbitrary loss" function to quantify the degradation of utility obtained by querying the LBS with the output of the mechanism instead of the real location of the user. The expected (*i.e.*, average) value of this function defines the expected loss of the mechanism. This value is a function of the specification of the mechanism and the prior probability distribution over possible locations of the user. There is a strong relationship between the expected loss of a mechanism and the imposed privacy constraints, which restricts the distinguishability between the points of $\mathcal{X}$. For instance, if we want two points to be entirely indistinguishable from each other, the output of the mechanism has to be independent of its input. This would lead to a maximization of the expected loss, thus rendering the mechanism useless. Thus to preserve some utility, the reported location should reflect (at least partially) the real location of the user, which necessitates a relaxed level of privacy (*i.e.*, a non-zero level of distinguishably between the points of $\mathcal{X}$).

Following the distinction made by the authors of [8] between *sporadic* and *continuous* location exposure, in this paper we will focus on the sporadic type. A typical example of LBS

corresponding to sporadic exposure is the search for nearby points of interest. In this situation, the locations reported by a user are sparsely distributed over time, and thus they can be reasonably assumed to be independent. In contrast in the continuous location exposure case, the locations reported are rather correlated due to their spatio-temporal closeness. These correlations can be exploit by the adversary to attack more successfully the location privacy of users. While we do not address this later case in this paper, recent work have started to address this issue such as the predictive mechanism [9] in the setting of $\epsilon$-geo-indistinguishability.

First, for the sake of simplicity and clarity, we focus our analysis on the specific model of $(D, \epsilon)$-location privacy. Later in Section 6, we extend our analysis and results to cover a broader range of constraints on the distinguishability between the points of $\mathcal{X}$. More precisely, we consider the general notion of $\ell$-privacy on an arbitrary set $\mathcal{X}$ of locations. In $\ell$-privacy, the distinguishability between two positions in $\mathcal{X}$ is quantified by a generic function $\ell$ depending on the distance between the given points. Afterwards, we show how our results can then be abstracted to also hold for $\ell$-privacy. According to the choice of $\ell$, $\ell$-privacy can be instantiated to various models of location privacy including $(D, \epsilon)$-location privacy and $\epsilon$-geo-indistinguishability.

Our contributions and well as the outline of the paper can be summarized as follows.

- After reviewing the related work (Section 2), we provide a probabilistic model for obfuscation mechanisms along with a generic utility metric for them (Section 3).
- We introduce the notion of $(D, \epsilon)$-location privacy and characterize its satisfiability for a mechanism in terms of the underlying characteristics of the mechanism and also the observer's prior and posterior knowledge (Section 3).
- We describe a specific class of mechanisms, which we refer to as "symmetric", obfuscating the location by adding noise to the real location of the user through a noise function (Section 4). We provide an analysis of such mechanisms with respect to location privacy and utility.
- Focusing on noise functions, we specify a subclass of them that we call as "circular", and prove that under certain conditions on $\mathcal{X}$, they are general enough to capture the same privacy guarantees and utility of any other non-circular function (Section 5).
- We generalize in Section 6 the constraints on the distinguishability to hold between arbitrary positions of $\mathcal{X}$ instead of adjacent ones. This generalization yields the notion $\ell$-privacy, which captures both $(D, \epsilon)$-location privacy and $\epsilon$-geo-indistinguishability [7]. We also extend our formal results to $\ell$-privacy.
- Finally, we compare our generic framework of $\ell$-privacy to other probabilistic notions of location privacy, namely the expected adversary's error [10] and $\epsilon$-geo-indistinguishability [7] with respect to their privacy guarantees and utility in Section 7 before concluding in Section 8.

## 2 Related Work

An early approach proposed for preserving the privacy of a user was to remove his identity or to replace it with pseudonyms [1]. However, this approach is bound to fail in the context of LBSs. Indeed, the user can be de-anonymized by the adversary (*e.g.*, the LBS provider)

from his reported location by correlating this data with background information such as household names [11] or employee-office correspondence [12]. As a consequence, while pseudonymizing a user is a first step, it is not sufficient to protect his privacy. Thus in addition, the request itself (*i.e.*, its location) should also be perturbed before being supplied to the LBS provider.

One type of approach, which proposed to "hide the location of the user inside a crowd", can be seen as the spatio-temporal equivalent of $k$-anonymity [13]. In the context of databases, the main guarantee provided by $k$-anonymity is that the record of one individual will be indistinguishable from $k-1$ others. This notion was adapted to protect the location of users by enabling the user to query an LBS using a spatial area called a "cloak" instead of his exact location. The system is built such that in this region there are at least $k-1$ individuals in addition to the user. Some of the early papers following this approach [3, 4] have also tried to quantify the trade-off between the offered privacy guarantees (quantified by $k$) and the resulting utility (quantified by the cloak size). For instance in [3], the authors introduced a basic model for location $k$-anonymity and proposed to rely on quadtrees to produce the smallest cloak while satisfying $k$-anonymity, while in [4] the model is extended such that each user can specify his desired privacy level.

The $k$-anonymity approach was also used for designing online distributed LBS. For example, the authors of [14] constructed a privacy-aware query processing framework called Casper, which processes the queries of users based on their cloaked (*i.e.*, anonymous) location. The major drawback of the spatio-temporal variant of $k$-anonymity is that it requires the use of a trusted third party playing the role of the anonymity server. This trusted third party has access to the locations of users and obfuscates them into cloaks when a user requests a service. Additionally the notion of $k$-anonymity was shown to be limited as it is vulnerable to the prior knowledge of the adversary about users [15].

Based on this criticism, some authors have proposed to quantify location privacy as the estimation error of the adversary when performing an inference attack [10, 8]. For instance, the main objective of the inference attack could be to deduce the true location of the user from the observed one or to re-identify the user based on the location disclosed. However, one of the drawbacks of this approach is that it needs to make strong assumptions on the knowledge available to the adversary to reason on the offered privacy level. For instance, in [10] the knowledge of the adversary is represented as a Markov model while in [8] the adversary is assumed to know the geographical distribution of users across different regions as well as their mobility patterns. Thus, while this approach is an important step towards the formalization of location privacy, the privacy guarantees offered are highly dependent on the prior knowledge of the adversary. In contrast, in this paper we aim at providing privacy guarantees that are *independent of the knowledge of the adversary*. We provide a detailed comparison of our proposal and related work quantifying location privacy with respect to the estimation error of the adversary in Section 7.1.

With respect to applying the differential privacy approach to location data, two other papers have been proposed [16, 17], in addition to geo-indistinguishability that we further detail in Section 7.2. The first paper [16] deals with the differentially private computation of the points of interest from a geographical database populated with the mobility traces of individuals based on a quadtree algorithm. The second paper [17] considers a database containing commuting patterns, in which each record corresponds to an individual and contains his origin and destination. Instead of publishing the original data, a synthetic dataset is generated mimicking the original data. This synthetic dataset is obtained by sampling from distributions learnt from the original data. While the authors showed that the original definition of differential privacy is too strong for this application as it guards

against very unlikely privacy breaches, they also demonstrated that their technique satisfies a weaker notion of differential privacy called *probabilistic differential privacy*.

# 3 A Framework for Location Privacy

In this section, we describe a probabilistic framework for analyzing the location privacy of the users. This framework consists of a model for the obfuscation mechanisms working on the locations of users as inputs, the characterization of the $(D, \epsilon)$-location privacy as a property offered by such mechanisms, and finally a metric for their utility.

## 3.1 Obfuscation Mechanisms for Locations

In the context of location data, we define an obfuscation mechanism as a *probabilistic function* $\mathcal{K}$ from a set $\mathcal{X} \subseteq \mathbb{E}^2$ of locations to the entire planar Euclidean space $\mathbb{E}^2$ [1]. The mechanism $\mathcal{K}$ takes as input the real location of a user and produces as output a location drawn at random from $\mathbb{E}^2$. We model this process by associating to every location $\boldsymbol{i} \in \mathcal{X}$ a (conditional) probability density function (pdf) $\mathcal{F}_{\boldsymbol{i}} : \mathbb{E}^2 \rightarrow \mathbb{R}^+$. More formally, given that the input of the mechanism is $\boldsymbol{i}$, the output is a continuous random variable (ranging on $\mathbb{E}^2$) whose pdf is $\mathcal{F}_{\boldsymbol{i}}$. We will refer to $\mathcal{F}_{\boldsymbol{i}}$ as the *randomization function* of $\mathcal{K}$ when the input is $\boldsymbol{i}$.

For allowing practical sampling of the outputs using randomization functions, we restrict these functions to be *bounded* and also continuous almost everywhere in $\mathbb{E}^2$. Precisely, within any bounded region of $\mathbb{E}^2$, a randomization function is assumed to be continuous everywhere except on a finite number of analytic curves[2].

**Definition 1.** (*Randomization function*) A randomization function $\mathcal{F} : \mathbb{E}^2 \rightarrow \mathbb{R}^+$ is a bounded probability density function on $\mathbb{E}^2$ such that inside any bounded region $S \subset \mathbb{E}^2$, the discontinuity of $\mathcal{F}$ is restricted to finitely many analytic curves.

Two examples of such functions are shown in Figure 1. In the first example, the discontinuity points form discrete circles having a common center while in the second example, the discontinuity points form a straight line. Modelling the output domain of a mechanism $\mathcal{K}$ by $\mathbb{E}^2$ does not cause any loss of generality even when the LBS provider accepts outputs in a different domain $\mathcal{Z}$ (*e.g.*, discrete latitude/longitude coordinates, block numbers or city names). In this situation, the output point $\boldsymbol{p}$ of $\mathcal{K}$ can be mapped to the appropriate element of $\mathcal{Z}$. As shown later, this mapping does not affect the location privacy if it is independent of the real location of the user (*e.g.*, if the mapping function is used by the LBS to map the reported location to a point of interest).

## 3.2 $(D, \epsilon)$-Location Privacy

Our notion of $(D, \epsilon)$-location privacy is an adaptation of the standard $\epsilon$-differential privacy [6] to LBSs. Differential privacy can be viewed as a property of mechanisms processing databases as secret information. More precisely, a mechanism that processes databases should provide a form of indistinguishability between every two *adjacent* databases, in which "adjacent" means that they differ only in a single record. In the context of location

[1] We use the notation $\mathbb{E}^2$ as it abstracts away from the underlying coordinate system unlike others (*e.g.*, $\mathbb{R}^2$ in which the Cartesian coordinate system is used).

[2] These curves have the property that their points are described by parametric functions that are "smooth" enough to be expressed by power series [18].

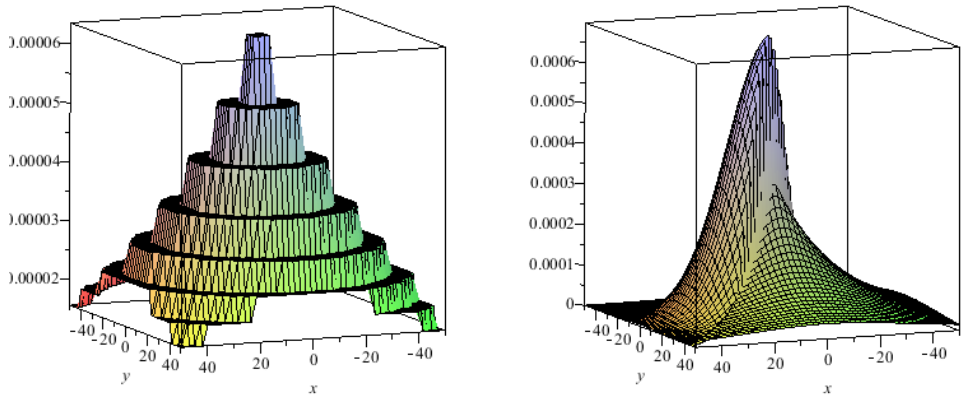



Figure 1: Examples of randomization functions.

privacy, we view the real location of the user as the secret to be protected, and accordingly (in the sense of differential privacy) we require a mechanism processing locations to ensure that every two "adjacent" locations are indistinguishable to some extent. We adapt the adjacency relation to this new context by saying that two locations $\boldsymbol{i}$ and $\boldsymbol{j}$ are adjacent to each other if the distance between them, written as $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j})$, is within a predefined proximity $D$ (*i.e.*, $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$).

We can now define $(D, \epsilon)$-location privacy for adjacent locations, in the same manner as the standard differential privacy is defined for adjacent databases. An obfuscation mechanism $\mathcal{K}$ satisfies $(D, \epsilon)$-location privacy if *every* two adjacent locations are indistinguishable (to a certain extent) from each other, when any output of $\mathcal{K}$ is observed. The adjacency relation is determined by the value of $D$ as mentioned previously, while the indistinguishability is quantified by the other parameter $\epsilon > 0$.

**Definition 2.** (*$(D, \epsilon)$-location privacy*) For a distance $D > 0$ and a real value $\epsilon > 0$, a mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfies $(D, \epsilon)$-location privacy if it holds for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ with $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$ that

$$P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\epsilon}\, P(\mathcal{K}(\boldsymbol{j}) \in S) \quad \forall S \subseteq \mathbb{E}^2.$$

According to the above definition, the probability of an observation $S$, given that the real location is $\boldsymbol{i}$ is within a multiplicative factor $e^{\epsilon}$ of the probability of the same observation given an adjacent location $\boldsymbol{j}$ (*i.e.*, within distance $D$ from $\boldsymbol{i}$). Basically, this prevents an adversary (*e.g.*, the LBS provider) observing the output of the mechanism from distinguishing the real location of a user from others situated within distance $D$.

## 3.3 Impact on the Adversary's Knowledge

Similarly to differential privacy, the definition of $(D, \epsilon)$-location privacy is described by the conditional probabilities of observations given inputs. Thus, the satisfiability of $(D, \epsilon)$-location privacy depends solely on the obfuscation mechanism itself, and abstracts away from the adversary's prior knowledge, which is usually modelled as a probability distribution on locations. We justify this abstraction from the prior knowledge using arguments

that are similarly used for the standard differential privacy [6] and geo-indistinguishability [7] as follows.

First, we emphasize that it is not possible for a privacy definition to guarantee the indistinguishability of the user's location under *any* prior knowledge while allowing a reasonable utility at the same time. In particular, the adversary's prior knowledge may enable him to infer the user's location from a fairly useful observation. For instance, consider a user located in some sparsely populated area (*e.g.*, the outskirt of a city containing a single restaurant). Then, an observation indicating that the user is positioned inside this area at the lunch time is enough for the adversary to guess his location. Given this constraint, our objective is not to protect the user's location against the adversary's prior knowledge, but rather we aim at restricting the impact of the output of the mechanism on such knowledge. More formally, we require the posterior probability distribution on locations to be relatively similar to the prior one. As shown by the following proposition, this demand is met by $(D, \epsilon)$-location privacy, which we characterize in terms of the adversary's prior and posterior knowledge (distributions) on any subset $\mathcal{I}$ of candidate locations of the user $\mathcal{X}$. For a location $\boldsymbol{i} \in \mathcal{I}$, we denote by $\pi_{\mathcal{I}}(\boldsymbol{i})$ the prior probability that $\boldsymbol{i}$ is the real location, and by $\pi_{\mathcal{I}}(\boldsymbol{i} \mid S)$ the posterior probability that the real location is $\boldsymbol{i}$ given that the mechanism output is in the region $S \subseteq \mathbb{E}^2$.

**Proposition 3** (Impact on the adversary's knowledge)**.** *Let $\mathcal{I}$ be any discrete set of locations. A mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfies $(D, \epsilon)$-location privacy if and only if it holds for all $\mathcal{I} \subseteq \mathcal{X}$ and all distributions $\pi_{\mathcal{I}}(.)$ on $\mathcal{I}$ that*

$$\pi_{\mathcal{I}}(\boldsymbol{i} \mid S)/\pi_{\mathcal{I}}(\boldsymbol{j} \mid S) \leq e^{\epsilon}\pi_{\mathcal{I}}(\boldsymbol{i})/\pi_{\mathcal{I}}(\boldsymbol{j}) \qquad \forall \boldsymbol{i}, \boldsymbol{j} \in \mathcal{I} : \mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D, \forall S \subseteq \mathbb{E}^2.$$

*Proof.* The condition of $(D, \epsilon)$-location privacy in Definition 2 can be written as $P(S \mid \boldsymbol{i}) \leq e^{\epsilon} P(S \mid \boldsymbol{j}), \forall S \subseteq \mathbb{E}^2, \forall \mathcal{I} \subseteq \mathcal{X}, \forall \boldsymbol{i}, \boldsymbol{j} \in \mathcal{I} : \mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$. The proof is completed by multiplying this inequality by $\pi_{\mathcal{I}}(\boldsymbol{i})\, \pi_{\mathcal{I}}(\boldsymbol{j})/P(S)$ in which $\pi_{\mathcal{I}}(.)$ is any prior distribution on $\mathcal{I}$. □

The above proposition means that the impact of the observation $S$ on the ratio of probabilities of two locations depends on the distance between them. In particular for adjacent locations $\boldsymbol{i}$ and $\boldsymbol{j}$ (which are at most $D$ apart), this ratio is multiplied by at most $e^{\epsilon}$. Thus, the mechanism itself does not substantially (subject to $\epsilon$) help the adversary to distinguish between them. However, for locations $\boldsymbol{i}'$ and $\boldsymbol{j}'$ that are further away from each other (*e.g.*, a restaurant in Paris and another one in London), the above ratio is allowed to be magnified by factor larger than $e^{\epsilon}$, thus allowing the observer to distinguish between them.

## 3.4 Role of Randomization Functions

Since any mechanism $\mathcal{K}$ is fully described by its randomization functions $\mathcal{F}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \mathcal{X}$, the satisfiability of $(D, \epsilon)$-location privacy for $\mathcal{K}$ can be characterized using these functions. This characterization is very useful as it provides an abstract, and yet simple, basis for analyzing such mechanisms.

**Theorem 4** (Characterization of $(D, \epsilon)$-location privacy)**.** *Let $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ be a mechanism with randomization functions $\mathcal{F}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \mathcal{X}$. This mechanism $\mathcal{K}$ satisfies $(D, \epsilon)$-location privacy if and only if for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ and all $\boldsymbol{p} \in \mathbb{E}^2$ such that $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$ and $\mathcal{F}_{\boldsymbol{i}}, \mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$, it holds that:*

$$\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\epsilon}\mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p}). \tag{1}$$

*Proof.* In $\iint_S \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p})\, d\lambda(\boldsymbol{p})$, we use $\lambda$ to denote the Lebesgue (area) measure defined on $\mathbb{E}^2$. We also use $\lambda(\boldsymbol{p})$ to emphasize that the integration is on the points $\boldsymbol{p}$ of $S$ (*cf.* [19, p102] and [20, p162]). Let $\mathcal{F}$ be a pdf on $\mathbb{E}^2$ that is continuous at $\boldsymbol{p} \in \mathbb{E}^2$. First, we prove that the average value of $\mathcal{F}$ around $\boldsymbol{p}$ converges to $\mathcal{F}(\boldsymbol{p})$. For any $\boldsymbol{p} \in \mathbb{E}^2$ and any $\delta > 0$, let $B_\delta(\boldsymbol{p}) \subset \mathbb{E}^2$ be the planar *ball* (neighborhood) centered at $\boldsymbol{p}$ and having radius $\delta$. We denote by $|B_\delta(\boldsymbol{p})|$ the area of $B_\delta(\boldsymbol{p})$. For a randomization (pdf) function $\mathcal{F}$ on $\mathbb{E}^2$, the function $\bar{\mathcal{F}}^{\boldsymbol{p}} : (0, \infty) \to \mathbb{R}^+$ of $\delta > 0$ is defined to be the *average value* of $\mathcal{F}$ in $B_\delta(\boldsymbol{p})$ as follows

$$\bar{\mathcal{F}}^{\boldsymbol{p}}(\delta) = 1/|B_\delta(\boldsymbol{p})| \iint_{B_\delta(\boldsymbol{p})} \mathcal{F}(\boldsymbol{p}')\, d\lambda(\boldsymbol{p}'), \qquad \forall \delta > 0.$$

Note that for all $\boldsymbol{p} \in \mathbb{E}^2, \delta > 0$ the above integral exists (*i.e.*, is finite) since $\mathcal{F}$ is bounded in $\mathbb{E}^2$. Since $\mathcal{F}$ is continuous at $\boldsymbol{p}$ then for every real $\sigma > 0$ there is a $\delta > 0$ such that $|\mathcal{F}(\boldsymbol{p}') - \mathcal{F}(\boldsymbol{p})| < \sigma, \forall \boldsymbol{p}' \in B_\delta(\boldsymbol{p})$. This also means that for all $\delta' < \delta$ we have $|\mathcal{F}(\boldsymbol{p}') - \mathcal{F}(\boldsymbol{p})| < \sigma, \forall \boldsymbol{p}' \in B_{\delta'}(\boldsymbol{p})$.

By the linearity of integrals (*cf.* [21, Proposition 4.2.5 ]), this inequality implies that $|\bar{\mathcal{F}}^{\boldsymbol{p}}(\delta') - \mathcal{F}(\boldsymbol{p})| < \sigma, \forall \delta' < \delta$. Thus, for all $\sigma > 0$ there is a $\delta > 0$ such that $|\bar{\mathcal{F}}^{\boldsymbol{p}}(\delta') - \mathcal{F}(\boldsymbol{p})| < \sigma$ for all $\delta' : 0 < \delta' < \delta$. This exactly defines the limit of $\bar{\mathcal{F}}^{\boldsymbol{p}}(\delta)$ when $\delta \to 0$. This means that for any pdf $\mathcal{F}$ that is continuous at $\boldsymbol{p}$, we have

$$\lim_{\delta \to 0} \bar{\mathcal{F}}^{\boldsymbol{p}}(\delta) = \mathcal{F}(\boldsymbol{p}). \tag{2}$$

We proceed by proving the theorem as follows. Suppose that a mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfies $(D, \epsilon)$-location privacy. Consider any $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ and any $\boldsymbol{p} \in \mathbb{E}^2$ in which $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$, and for which both randomization functions $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$. Applying Definition 2 to any ball $B_\delta(\boldsymbol{p})$, with radius $\delta > 0$, yields $\bar{\mathcal{F}}_{\boldsymbol{i}}^{\boldsymbol{p}}(\delta) \leq e^\epsilon\, \bar{\mathcal{F}}_{\boldsymbol{j}}^{\boldsymbol{p}}(\delta)$. Considering the limits in this inequality when $\delta \to 0$, and substituting these limits using Equation (2) yield Inequality (1).

Conversely, assume that Inequality (1) holds for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ and $\boldsymbol{p} \in \mathbb{E}^2$ in which $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$ and both $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$. Since $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous almost everywhere according to Definition 1, Inequality (1) holds also almost everywhere in $\mathbb{E}^2$. Thus, it holds for any region $S \subseteq \mathbb{E}^2$ by [21, Proposition 4.3.7 ] that $\iint_S \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p})\, d\lambda(\boldsymbol{p}) \leq \iint_S e^\epsilon\, \mathcal{F}_j(\boldsymbol{p})\, d\lambda(\boldsymbol{p})$, the condition of Definition 2. □

Note that this theorem does not restrict the values of the randomization functions at their discontinuity points. This directly follows from the fact that these functions (by Definition 1) are continuous almost everywhere in $\mathbb{E}^2$. In particular, the values of $\mathcal{F}_{\boldsymbol{i}}$ at its discontinuity points do not affect its integral over any region $S \subseteq \mathbb{E}^2$ when evaluating the probability $P(\mathcal{K}(\boldsymbol{i}) \in S)$ appearing in Definition 2. Thus, satisfying $(D, \epsilon)$-location privacy is independent of the values of randomization functions at their discontinuities. We remark also that If the functions $\mathcal{F}_{\boldsymbol{i}}$ are the same for all input locations, then $(D, \epsilon)$-location privacy is trivially satisfied. However, such a mechanism would be completely useless, as its output would be always drawn from the same pdf independently of the input, and thus carries no information about the location of the user.

## 3.5 Impact of the Post-Processing on Privacy

If a mechanism satisfies $(D, \epsilon)$-location privacy, then any post-processing of the output $\boldsymbol{p}$ of such a mechanism does not leak additional information about the location of the user. More precisely, assume that the output of $\mathcal{K}$ is processed by a mapping function $\mathcal{M}$ to produce

an output $z$ in another domain $\mathcal{Z}$ (*e.g.*, latitude/longitude coordinate system, points of interest, city names, ...). In this situation, given $z$ (and $\boldsymbol{p}$), the adversary is still unable (relative to $\epsilon$) to distinguish the location of a user from adjacent ones provided that this post-processing is done independently of the original location of the user. We formalize this observation by the following proposition in which $(\mathcal{M} \circ \mathcal{K})$ denotes the composition of $\mathcal{K}$ and $\mathcal{M}$.

**Proposition 5** (Post-processing preserves privacy)**.** *Consider a mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfying $(D, \epsilon)$-location privacy. Let also $\mathcal{M} : \mathbb{E}^2 \to \mathcal{Z}$ be a probabilistic mapping function in which $\mathcal{Z}$ is an arbitrary domain. Then, it holds for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ with $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$ that*

$$P\left((\mathcal{M} \circ \mathcal{K})(\boldsymbol{i}) \in Z\right) \leq e^{\epsilon}\, P\left((\mathcal{M} \circ \mathcal{K})(\boldsymbol{j}) \in Z\right) \quad \forall Z \subseteq \mathcal{Z}.$$

*Proof.* For any $\boldsymbol{i} \in \mathcal{X}$, $Z \subseteq \mathcal{Z}$, the probability $P((\mathcal{M} \circ \mathcal{K})(\boldsymbol{i}) \in Z)$ is evaluated by considering every output $\boldsymbol{p}$ of $\mathcal{K}(\boldsymbol{i})$ which is mapped by $\mathcal{M}$ to some element in $Z$ with probability $P(\mathcal{M}(\boldsymbol{p}) \in Z)$. This means that $P((\mathcal{M} \circ \mathcal{K})(\boldsymbol{i}) \in Z) = \iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p})\, P(\mathcal{M}(\boldsymbol{p}) \in Z)\, d\lambda(\boldsymbol{p})$. Consider now another point $\boldsymbol{j} \in \mathcal{X}$ in which $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$. Since $\mathcal{K}$ satisfies $(D, \epsilon)$-location privacy, it holds by Theorem 4 that $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\epsilon}\, \mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p})$ almost everywhere in $\mathbb{E}^2$. The proof is then completed by multiplying both sides of this inequality by $P(\mathcal{M}(\boldsymbol{p}) \in Z)$, and then integrating them over $\mathbb{E}^2$. □

In practice, the post-processing applied on the output of the mechanism $\mathcal{K}$ usually depends on the target LBS. For example, an LBS providing information about nearby restaurants will expect the latitude/longitude coordinates of the location, while a weather forecasting LBS may require only the city of the user.

## 3.6 Utility Model

While a mechanism is required to satisfy the location privacy for the user, it should also produce outputs that remains useful in terms of the LBS used. To quantify the utility of the mechanism to the user, we rely on the notion of *loss functions*. In a nutshell, the loss function measures the loss incurred by reporting an obfuscated location instead of the real one. Specifically, we model a loss function as a mapping $\mathcal{L} : [0, \infty) \to [0, \infty)$ taking as its argument the distance between the input and output locations of the mechanism, and evaluating to a real number that quantifies the loss due to the obfuscation. For instance, the loss can be proportional to the distance as $\mathcal{L}(d) = d$, or possibly increasing "faster" with the distance as $\mathcal{L}(d) = d^n$ for some $n > 1$.

Using a given loss function, we quantify the utility of a mechanism for a user as the expected loss value (simply called the *expected loss*). Since the output of the mechanism depends (probabilistically) on the location of the user, the expected loss depends on the (prior) probability distribution $\pi$ of the user on his arbitrary points of interests $\mathcal{I} \subseteq \mathcal{X}$ [3].

This distribution can be estimated by simply measuring the user's frequency of visiting each point $\boldsymbol{i} \in \mathcal{I}$. The expected loss can be formulated as follows.

**Definition 6.** (*Expected loss of a mechanism*) Consider a mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ with randomization functions $\mathcal{F}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \mathcal{X}$. Given a loss function $\mathcal{L} : [0, \infty) \to [0, \infty)$ and a prior

[3] The points of interests $\mathcal{I}$ are typically assumed to be discrete. However, our analysis remains the same even if $\mathcal{I}$ is assumed to be a continuous sub-region of $\mathcal{X}$, in which the prior is replaced by a pdf and the summation in Definition 6 is replaced by an integral.

$\pi$ on an arbitrary set of locations $\mathcal{I} \subseteq \mathcal{X}$, the *expected loss* of $\mathcal{K}$ is defined as

$$\Psi(\mathcal{K}, \mathcal{L}, \pi) = \mathbf{E}(\mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{p}))) = \sum_{\boldsymbol{i} \in \mathcal{I}} \pi(\boldsymbol{i}) \left( \iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p})\, \mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{p}))\, d\lambda(\boldsymbol{p}) \right).$$

The above quantification is similar to the approach taken in the analysis of standard differential privacy [22, 23, 24, 25]. However in our context, the user knows the secret (*i.e.*, his real location), in contrast to standard differential privacy in which the secret (here the database) is hidden from the user who only observes the output of the mechanism. This difference explains why in our paper a user does not need to guess the secret by *remapping* the output of the mechanism as done in the context of differential privacy.

The utility measure $\Psi$ of a mechanism (for a user) is also similar to the "expected quality loss" adopted in [26] as both of them evaluate the expected value of a user-defined loss function. However $\Psi$ requires that the outputs of the mechanism are sampled from the continuous domain $\mathbb{E}^2$, and thus its evaluation involves an integral on this domain.

By Definition 6, the expected loss of the mechanism depends generally on the prior $\pi$, which may differ from one user to another. In the next section, we focus on a specific class of mechanisms, called *symmetric mechanisms*, for which we demonstrate that their expected loss is independent of the user's prior. This characteristic enables us to compare the utilities of different mechanisms regardless of the priors of users.

# 4 Symmetric Mechanisms

In this section, we investigate the obfuscation of the user's location by the *addition* of random noise that is drawn independently of the input, through a special class of mechanisms that we call as "symmetric mechanisms". Before presenting these mechanisms in terms of their underlying randomization functions, we first introduce the notion of *noise vectors*.

In our modelling, the inputs and outputs of a mechanism are points in the Euclidean space $\mathbb{E}^2$. Therefore, an output point $\boldsymbol{p}$ can be viewed as the sum of the input point $\boldsymbol{i}$ of the mechanism and an Euclidean vector $\vec{\mu}$ called a noise vector (*i.e.*, $\boldsymbol{p} = \boldsymbol{i} + \vec{\mu}$).

In the rest of the paper, we consider a fixed point in $\mathbb{E}^2$, called as the *origin*, and denoted by $\boldsymbol{o}$. Relative to the origin $\boldsymbol{o}$, every vector $\vec{\mu}$ can be seen as the "position vector" of some point $\boldsymbol{u}$ (*i.e.*, $\vec{\mu}$ is the displacement from $\boldsymbol{o}$ to $\boldsymbol{u}$). This establishes a one-to-one correspondence between noise vectors and points of $\mathbb{E}^2$, which means that each vector $\vec{\mu}$ is the position vector of a unique point denoted by $\mathbf{pnt}(\vec{\mu})$. Conversely, each point $\boldsymbol{u}$ has a unique position vector denoted by $\mathbf{vec}(\boldsymbol{u})$[4].

With respect to noise vectors, we say that a mechanism is "symmetric" if at every run, the computation of the noise vector is independent of the input location. Using the correspondence with the points of $\mathbb{E}^2$, the computation of the noise vector can be modeled by drawing its end point from $\mathbb{E}^2$ according to some pdf $\mathcal{F}$. Therefore, the symmetry of a mechanism means that $\mathcal{F}$ is fixed (*i.e.*, independent of the input). This property can be formulated in terms of the randomization functions $\mathcal{F}_{\boldsymbol{i}}$ of $\mathcal{K}$ as follows.

**Definition 7.** (*Symmetric mechanisms*) A mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ is said to be *symmetric* if there is a fixed randomization function $\mathcal{F}$ such that:

$$\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) = \mathcal{F}\left(\, \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i}) \,\right) \qquad \forall \boldsymbol{i} \in \mathcal{X}, \forall \boldsymbol{p} \in \mathbb{E}^2. \tag{3}$$

[4]From the arithmetics of vectors, it holds for any point $\boldsymbol{p}$ and any vector $\vec{\mu}$ that $\mathbf{pnt}(\mathbf{vec}(\boldsymbol{p})) = \boldsymbol{p}$ and $\mathbf{vec}(\mathbf{pnt}(\vec{\mu})) = \vec{\mu}$. Moreover, for any pair of points $\boldsymbol{p}_1, \boldsymbol{p}_2$, we have $\mathbf{d}(\boldsymbol{p}_1, \boldsymbol{p}_2) = |\boldsymbol{p}_1 - \boldsymbol{p}_2| = |\mathbf{vec}(\boldsymbol{p}_1) - \mathbf{vec}(\boldsymbol{p}_2)|$ and $\boldsymbol{p}_1 + \mathbf{vec}(\boldsymbol{p}_2) = \mathbf{vec}(\boldsymbol{p}_1) + \boldsymbol{p}_2$.

This definition states that drawing an output point $\boldsymbol{p}$ when the input is $\boldsymbol{i}$ is equivalent to drawing the difference between them according to a fixed pdf $\mathcal{F}$. Since this difference corresponds to the added noise vector, we refer to $\mathcal{F}$ as the *noise function* of the mechanism. This noise function is exactly the randomization function of $\mathcal{K}$ when the input is the origin point $\boldsymbol{o}$ (*i.e.*, $\mathcal{F} = \mathcal{F}_{\boldsymbol{o}}$). Consequently, noise functions are essentially randomization functions (*cf.* Definition 1).

We call such a mechanism "symmetric" because it exhibits a *translational symmetry*, which means that the mechanism is invariant under any translation on $\mathbb{E}^2$. In fact, by Equation (3), the probability density at a point $\boldsymbol{p}$ when the input is $\boldsymbol{i}$ is the same if these points are replaced by (or translated to) $\boldsymbol{p} + \vec{\tau}$ and $\boldsymbol{i} + \vec{\tau}$ respectively, in which $\vec{\tau}$ is any vector (*i.e.*, $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) = \mathcal{F}_{\boldsymbol{i}+\vec{\tau}}(\boldsymbol{p} + \vec{\tau})$). A consequence of this symmetry is that the properties of the noise function $\mathcal{F}$ are translated to all randomization functions of the mechanism. In particular, the following lemma links the continuity of the randomization functions to $\mathcal{F}$.

**Lemma 8** (Continuity)**.** *For a symmetric mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ with the noise function $\mathcal{F}$, it holds for all $\boldsymbol{i} \in \mathcal{X}, \forall \boldsymbol{p} \in \mathbb{E}^2$ that $\mathcal{F}_{\boldsymbol{i}}$ is continuous at $\boldsymbol{p}$ if and only if $\mathcal{F}$ is continuous at $\mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$.*

*Proof.* A function $\mathcal{H} : \mathbb{E}^2 \to \mathbb{R}^+$ is continuous at a point $\boldsymbol{x} \in \mathbb{E}^2$ if and only if for every $\sigma > 0$, there exists $\delta > 0$, such that for all $\boldsymbol{x}' \in \mathbb{E}^2$ with $\mathbf{d}(\boldsymbol{x}, \boldsymbol{x}') < \delta$, it holds that $|\mathcal{H}(\boldsymbol{x}) - \mathcal{H}(\boldsymbol{x}')| < \sigma$. Using this definition of *continuity*, we prove the lemma as follows. For any $\boldsymbol{i} \in \mathcal{X}, \boldsymbol{p} \in \mathbb{E}^2$ and real $\sigma > 0$, let $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$ (*i.e.*, $\boldsymbol{p} = \boldsymbol{i} + \mathbf{vec}(\boldsymbol{u}) = \mathbf{vec}(\boldsymbol{i}) + \boldsymbol{u}$). We show that the following two statements are equivalent for a symmetric mechanism having the noise function $\mathcal{F}$:

1. $\exists \delta > 0$ such that for all $\boldsymbol{p}' \in \mathbb{E}^2$ in which $\mathbf{d}(\boldsymbol{p}, \boldsymbol{p}') < \delta$, it holds $\left|\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) - \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}')\right| < \sigma$.
2. $\exists \delta' > 0$ such that for all $\boldsymbol{u}' \in \mathbb{E}^2$ in which $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') < \delta'$, it holds $\left|\mathcal{F}(\boldsymbol{u}) - \mathcal{F}(\boldsymbol{u}')\right| < \sigma$.

First, we demonstrate that Statement 1 implies Statement 2. Set $\delta' = \delta$. Now for any point $\boldsymbol{u}'$ satisfying $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') < \delta'$ let $\boldsymbol{p}' = \boldsymbol{i} + \mathbf{vec}(\boldsymbol{u}')$. For this point $\boldsymbol{p}'$, we have $\mathbf{d}(\boldsymbol{p}, \boldsymbol{p}') = |(\boldsymbol{i} + \mathbf{vec}(\boldsymbol{u})) - (\boldsymbol{i} + \mathbf{vec}(\boldsymbol{u}'))| = \mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') < \delta' = \delta$. Thus, by Statement 1, it holds that $\left|\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) - \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}')\right| < \sigma$. Therefore, by Equation (3), it follows that $\left|\mathcal{F}(\boldsymbol{u}) - \mathcal{F}(\boldsymbol{u}')\right| < \sigma$. Repeating the same argument for all points $\boldsymbol{u}'$ in which $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') < \delta'$, we obtain Statement 2.

Similarly, Statement 1 is implied from Statement 2. Assume that 2 holds and set $\delta = \delta'$. For any point $\boldsymbol{p}'$ satisfying $\mathbf{d}(\boldsymbol{p}, \boldsymbol{p}') < \delta$ let $\boldsymbol{u}' = \mathbf{pnt}(\boldsymbol{p}' - \boldsymbol{i})$, i.e. $\boldsymbol{u}' = \boldsymbol{p}' - \mathbf{vec}(\boldsymbol{i})$. For this specific point, we have $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') = |(\boldsymbol{p} - \mathbf{vec}(\boldsymbol{i})) - (\boldsymbol{p}' - \mathbf{vec}(\boldsymbol{i}))| = \mathbf{d}(\boldsymbol{p}, \boldsymbol{p}') < \delta = \delta'$. Thus, by Statement 2, it holds that $\left|\mathcal{F}(\boldsymbol{u}) - \mathcal{F}(\boldsymbol{u}')\right| < \sigma$, which implies by Equation (3) that $\left|\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) - \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}')\right| < \sigma$. Repeating the same argument for all points $\boldsymbol{p}'$ in which $\mathbf{d}(\boldsymbol{p}, \boldsymbol{p}') < \delta$, we obtain Statement 1. □

To summarize, a symmetric mechanism obfuscates the input location by adding a noise vector $\vec{\mu}$ to it. This noise vector is the position vector of a point drawn according to the noise function $\mathcal{F}$. Thus, the characteristics of a symmetric mechanism are determined entirely by its noise function $\mathcal{F}$.

## 4.1 ($D$, $\epsilon$)-Location Privacy of Symmetric Mechanisms

In the following, we identify the sufficient and necessary condition for a symmetric mechanism to satisfy ($D$, $\epsilon$)-location privacy. Using Definition 7, we translate the characterization

of ($D$, $\epsilon$)-location privacy (stated in Theorem 4 for generic mechanisms) to a condition applied to the noise function $\mathcal{F}$ of the symmetric mechanism. Intuitively, this condition of privacy depends also on the domain $\mathcal{X}$ on which the mechanism is applied. More precisely, we find that this condition depends on the set $\mathcal{V}_\mathcal{X} = \{\boldsymbol{j} - \boldsymbol{i} : \boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}\}$ (*i.e.*, the set of all displacement vectors between the points of $\mathcal{X}$). Using this set, we can phrase the required conditions of privacy in the following theorem.

**Theorem 9** (($D$, $\epsilon$)-location private symmetric mechanisms)**.** *A symmetric mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ with a noise function $\mathcal{F} : \mathbb{E}^2 \to \mathbb{R}^+$ satisfies ($D$, $\epsilon$)-location privacy if and only if for all points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous, $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_\mathcal{X}$, and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) \leq D$, it holds that $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{v})$.*

*Proof.* We prove that for a symmetric mechanism having the noise function $\mathcal{F}$, the condition of ($D$, $\epsilon$)-location privacy in Theorem 4 is equivalent to the condition stated by Theorem 9 (*i.e.*, we prove that the following statements are equivalent).

1. $\forall \boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \forall \boldsymbol{p} \in \mathbb{E}^2$ such that $\mathcal{F}_\boldsymbol{i}, \mathcal{F}_\boldsymbol{j}$ are continuous at $\boldsymbol{p}$ and $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$, it holds that $\mathcal{F}_\boldsymbol{i}(\boldsymbol{p}) \leq e^\epsilon \mathcal{F}_\boldsymbol{j}(\boldsymbol{p})$.
2. $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ such that $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_\mathcal{X}$, $\mathcal{F}$ is continuous at $\boldsymbol{u}, \boldsymbol{v}$; and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) \leq D$, it holds that $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{v})$.

We assume that Statement 1 holds and show that Statement 2 follows. Consider any pair of points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous, $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_\mathcal{X}$, and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) \leq D$. We show that it must hold for them that $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{v})$. Let $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \boldsymbol{p} \in \mathbb{E}^2$ be any points satisfying $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$ and $\boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$. These points exist for $\boldsymbol{u}$ and $\boldsymbol{v}$ in the following manner. Since $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_\mathcal{X}$, there exists $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ in which $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{i} - \boldsymbol{j}$. Now let $\boldsymbol{p} = \mathbf{vec}(\boldsymbol{i}) + \boldsymbol{u} = \mathbf{vec}(\boldsymbol{j}) + \boldsymbol{v}$. Then it is easy to see that $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{p}$ satisfy that $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$ and $\boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$.
Since $\mathcal{F}$ is continuous at $\boldsymbol{u}, \boldsymbol{v}$, it must hold (by Lemma 8) that $\mathcal{F}_\boldsymbol{i}$ and $\mathcal{F}_\boldsymbol{j}$ are continuous at the point $\boldsymbol{p}$. In addition, it holds that $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) = |(\boldsymbol{p} - \mathbf{vec}(\boldsymbol{u})) - (\boldsymbol{p} - \mathbf{vec}(\boldsymbol{v}))| = \mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) \leq D$. From the assumption that Statement 1 holds, it follows that $\mathcal{F}_\boldsymbol{i}(\boldsymbol{p}) \leq e^\epsilon \mathcal{F}_\boldsymbol{j}(\boldsymbol{p})$, which implies by Definition 7 that $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{v})$. By repeating the same argument for all pairs $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) \leq D$, we conclude that Statement 2 holds.
Now we assume that Statement 2 holds and show that Statement 1 is implied. Consider any triplet of points $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \boldsymbol{p} \in \mathbb{E}^2$ in which $\mathcal{F}_\boldsymbol{i}$ and $\mathcal{F}_\boldsymbol{j}$ are both continuous at $\boldsymbol{p}$ and $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$. We demonstrate that they must satisfy $\mathcal{F}_\boldsymbol{i}(\boldsymbol{p}) \leq e^\epsilon \mathcal{F}_\boldsymbol{j}(\boldsymbol{p})$. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ be any points satisfying $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i}), \boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$. Such points always exist (and are unique) for $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{p}$. Now since $\mathcal{F}_\boldsymbol{i}$ and $\mathcal{F}_\boldsymbol{j}$ are continuous at $\boldsymbol{p}$, it must hold (by Lemma 8) that $\mathcal{F}$ is continuous at $\boldsymbol{u}$ and $\boldsymbol{v}$. It also holds that $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{i} - \boldsymbol{j} \in \mathcal{V}_\mathcal{X}$ and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) = \mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$. By the assumption that Statement 2 holds, we have for $\boldsymbol{u}, \boldsymbol{v}$ that $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{v})$ which implies, by Definition 7, that $\mathcal{F}_\boldsymbol{i}(\boldsymbol{p}) \leq e^\epsilon \mathcal{F}_\boldsymbol{j}(\boldsymbol{p})$. Repeating the same argument for all triplets $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{p} \in \mathbb{E}^2$ in which $\mathcal{F}_\boldsymbol{i}$ and $\mathcal{F}_\boldsymbol{j}$ are continuous at $\boldsymbol{p}$ and $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$, we obtain Statement 1. □

Satisfying ($D$, $\epsilon$)-location privacy is independent of the values of $\mathcal{F}$ at its discontinuity points in $\mathbb{E}^2$. By Definition 7 and Lemma 8, these values match those of the randomization functions $\mathcal{F}_\boldsymbol{i}$ at their respective discontinuities. Therefore, they do not impact the satisfiability of ($D$, $\epsilon$)-location privacy as discussed earlier in Section 3. In addition, satisfying ($D$, $\epsilon$)-location privacy requires that $\mathcal{F}$ is strictly non-zero at all points of $\mathbb{E}^2$ at which $\mathcal{F}$ is continuous. This requirement can be justified informally in the following manner. Assume that $\mathcal{F}$ is known to be continuous at the point $\boldsymbol{u}$ and $\mathcal{F}(\boldsymbol{u}) = 0$, which means that noise

vectors are unlikely to take values close to $\mathbf{vec}(\boldsymbol{u})$. Consider now an adversary observing a point $\boldsymbol{p}$ as the output of the mechanism. From his knowledge about $\mathcal{F}$, the adversary can rule out the possibility that the user is in a small region (*i.e.*, neighborhood) $S$ surrounding the point $\boldsymbol{p} - \mathbf{vec}(\boldsymbol{u})$ (*i.e.*, distinguish the points of $S$ from their adjacent points), which violates the ($D$, $\epsilon$)-location privacy of the user.

## 4.2 Expected Loss for Symmetric Mechanisms

Given a loss function $\mathcal{L}$, we can express the expected loss $\Psi$ of a symmetric mechanism using Definition 6. Indeed, since this mechanism is fully characterized by its noise function $\mathcal{F}$, the expected loss depends also on $\mathcal{F}$. However, unlike non-symmetric mechanisms, the expected loss for this type of mechanism is independent of the prior probability distribution of the user on his points of interest. This result follows directly from the fact that the noise vector is drawn according to the noise function $\mathcal{F}$ regardless of the input of the mechanism.

**Proposition 10** (Expected loss of a symmetric mechanism)**.** *Let $\mathcal{K}$ be a symmetric mechanism with a noise function $\mathcal{F}$. Given a loss function $\mathcal{L}$ and any prior $\pi$, the* expected loss *of $\mathcal{K}$ (and also $\mathcal{F}$) with respect to $\mathcal{L}$ is given by:*

$$\Psi(\mathcal{K}, \mathcal{L}) = \iint_{\mathbb{E}^2} \mathcal{F}(\boldsymbol{u})\, \mathcal{L}\left(\mathbf{d}(\boldsymbol{o}, \boldsymbol{u})\right)\, d\lambda(\boldsymbol{u}).$$

*Proof.* If the input of the mechanism is $\boldsymbol{i} \in \mathcal{X}$, the expected value of the loss (given $\boldsymbol{i}$) is evaluated by the integral $\iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{y})\, \mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{y}))\, d\lambda(\boldsymbol{y})$. We proceed by showing that the value of this integral is independent of $\boldsymbol{i}$, using the technique of *transformations* (*cf.* [19, sec. 39]) as follows.

For a given point $\boldsymbol{i} \in \mathcal{X}$, we define the transformation $T_i : \mathbb{E}^2 \to \mathbb{E}^2$ as $T_i(\boldsymbol{x}) = \boldsymbol{i} + \mathbf{vec}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{E}^2$. Let also the function $g : \mathbb{E}^2 \to [0, \infty)$ be such that $g(\boldsymbol{y}) = \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{y})\, \mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{y}))$ for all $\boldsymbol{y} \in \mathbb{E}^2$. By [19, Theorem 39.C], we have that

$$\iint_{\mathbb{E}^2} g(\boldsymbol{y})\, d\lambda T_i^{-1}(\boldsymbol{y}) = \iint_{\mathbb{E}^2} g(T_i(\boldsymbol{x}))\, d\lambda(\boldsymbol{x}), \tag{4}$$

in which the measure $\lambda T_i^{-1}$ on the subsets of $\mathbb{E}^2$ is defined as $\lambda T_i^{-1}(S) = \lambda(T_i^{-1}(S))$ for every measurable subset $S \subseteq \mathbb{E}^2$. Note that $\lambda T_i^{-1}(S)$ is exactly the area of the region $T_i^{-1}(S)$, which is mapped to $S$ by $T_i$ and that the areas of $S$ and $T_i^{-1}(S)$ are equal due to the definition of $T_i$. Thus, we have that $\lambda T_i^{-1}(S) = \lambda(S)$ for every $\lambda T_i^{-1}$-measurable set $S \subseteq \mathbb{E}^2$. By substituting $g, T_i$, and $\lambda T_i^{-1}$ in Equation (4), we get

$$\iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{y})\, \mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{y}))\, d\lambda(\boldsymbol{y}) = \iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{i} + \mathbf{vec}(\boldsymbol{x}))\, \mathcal{L}(\mathbf{d}(\boldsymbol{i}, \boldsymbol{i} + \mathbf{vec}(\boldsymbol{x})))\, d\lambda(\boldsymbol{x}).$$

By Definition 7, and the fact that $\mathbf{pnt}(\mathbf{vec}(\boldsymbol{x})) = \boldsymbol{x}$ and $\mathbf{d}(\boldsymbol{i}, \boldsymbol{i} + \mathbf{vec}(\boldsymbol{x}))) = \mathbf{d}(\boldsymbol{o}, \boldsymbol{x})$ we can simplify the latter integral to $\iint_{\mathbb{E}^2} \mathcal{F}(\boldsymbol{x})\, \mathcal{L}(\mathbf{d}(\boldsymbol{o}, \boldsymbol{x}))\, d\lambda(\boldsymbol{x})$.

The resulting integral is independent of $\boldsymbol{i}$. The proof is completed by substituting this integral in Definition 6 and using the fact that $\sum_{i \in \mathcal{I}} \pi_{\boldsymbol{i}} = 1$. □

The above result holds both if the prior $\pi$ is a probability distribution on a discrete set $\mathcal{I}$ of points or a probability density function on $\mathbb{E}^2$. Thus, the utility of symmetric mechanisms

can be analyzed without taking into account the prior of the user. Note that the expression of $\Psi$ can also be interpreted in terms of noise vectors. Indeed, since $\mathbf{d}(\boldsymbol{o}, \boldsymbol{u})$ is exactly the magnitude of the random noise vector, $\Psi$ is actually the expected value of the loss function $\mathcal{L}$ computed from the magnitudes of noise vectors. For example if $\mathcal{L}(d) = d$, then $\Psi$ is exactly the average noise magnitude, which is also the average distance between the real and reported locations.

To summarize, both the location privacy guarantee and the expected loss of a symmetric mechanism are independent from the prior knowledge (of the adversary and the user) as they depend only on the noise function of the mechanism. In addition, it turns out that under certain assumptions about the domain $\mathcal{X}$, a special family of noise functions called "circular", which we describe in the next section, is sufficient to capture all values of expected loss that can be achieved under given privacy constraints.

# 5 Circular Noise Functions

Taking advantage of the assumption that the loss function depends only on the magnitude of the associated noise vector, we consider a family of noise functions in which each member possesses a special uniformity feature. In particular, since noise vectors having the same magnitude share the same loss value, there is no advantage of assigning to them different probability densities. These vectors correspond to the points of a circle centered on the origin $\boldsymbol{o}$. Thus, in this section, we study the idea of assigning to these points the same probability density, yielding a special class of noise functions, which we coin as *circular*.

**Definition 11.** (*Circular noise function*) A noise function $\mathcal{F} : \mathbb{E}^2 \to \mathbb{R}^+$ is said to be *circular* if there is a function $\mathcal{R} : [0, \infty) \to \mathbb{R}^+$, called the *radial* of $\mathcal{F}$, such that for all $\boldsymbol{u} \in \mathbb{E}^2$ it holds:

$$\mathcal{F}(\boldsymbol{u}) = \mathcal{R}(\mathbf{d}(\boldsymbol{o}, \boldsymbol{u})).$$

The above definition states that a circular noise function assigns one probability density value to all points at the same distance from $\boldsymbol{o}$. This value is determined by a certain radial $\mathcal{R}$. A circular noise function is fully specified by its radial $\mathcal{R}$. Thus, the analysis of circular noise functions can be reduced to the analysis of their associated, and simpler, radials whose variable is a radius $r \in [0, \infty)$. Hereafter, we denote by $\mathcal{F}_{\mathcal{R}}$ the circular noise function whose radial is $\mathcal{R}$.

Several properties of $\mathcal{R}$ are inherited from its noise function $\mathcal{F}_{\mathcal{R}}$. In particular, $\mathcal{R}$ (as well as $\mathcal{F}_{\mathcal{R}}$) is bounded and continuous almost everywhere on its domain. More precisely, within any bounded subinterval of $[0, \infty)$, the radial $\mathcal{R}$ is discontinuous only at finitely many points. These points correspond to circles in the domain $\mathbb{E}^2$ of $\mathcal{F}_{\mathcal{R}}$. Furthermore, since $\mathcal{F}_{\mathcal{R}}$ is a pdf, it must satisfy the *total probability law*, stating that its integral over its domain $\mathbb{E}^2$ evaluates to 1. If we express the integral in terms of the radial $\mathcal{R}$, we obtain

$$\int_0^\infty \mathcal{R}(r)\, 2\pi r\, dr = 1. \tag{5}$$

## 5.1 ($D$, $\epsilon$)-Location Privacy and Utility

We recall that the privacy characteristics (and utility) of a symmetric mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ are determined by its underlying noise function. In the following we express the sufficient and necessary conditions on circular noise functions to satisfy ($D$, $\epsilon$)-location privacy.

These conditions do not depend only on the privacy parameters but also on the domain $\mathcal{X}$ on which the mechanism is applied. More precisely, these conditions are related to $\epsilon$ as well as to the set $\Omega_{\mathcal{X},D}$ defined as

$$\Omega_{\mathcal{X},D} = \{(|\boldsymbol{u}|, |\boldsymbol{u}'|) : \boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2, \boldsymbol{u}' - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}, |\boldsymbol{u}' - \boldsymbol{u}| \leq D\}.$$

To understand the intuition behind $\Omega_{\mathcal{X},D}$, recall that every two points $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ having $\boldsymbol{u}' - \boldsymbol{u} \in \mathcal{V}$ and $|\boldsymbol{u}' - \boldsymbol{u}| \leq D$ should be indistinguishable (relative to $\epsilon$) according to Theorem 9. This means that their respective distances $(|\boldsymbol{u}|, |\boldsymbol{u}'|)$ from the origin $\boldsymbol{o}$ have also to be made indistinguishable by the circular noise function $\mathcal{F}_{\mathcal{R}}$. Based on this intuition, we state the required conditions of the privacy for $\mathcal{F}_{\mathcal{R}}$ in terms of its radial $\mathcal{R}$ and also $\Omega_{\mathcal{X},D}$ as follows.

**Corollary 12** (($D$, $\epsilon$)-location privacy of circular noise functions)**.** *A circular noise function $\mathcal{F}_{\mathcal{R}}$ having a radial $\mathcal{R}$ satisfies ($D$, $\epsilon$)-location privacy for a domain $\mathcal{X}$ if and only if for all $(r, r') \in \Omega_{\mathcal{X},D}$ such that $\mathcal{R}$ is continuous at $r$ and $r'$, we have: $\mathcal{R}(r) \leq e^{\epsilon}\, \mathcal{R}(r')$.*

*Proof.* Assume that $\mathcal{F}_{\mathcal{R}}$ satisfies ($D$, $\epsilon$)-location privacy for a domain $\mathcal{X}$. In the following, we show that under this assumption the condition in the corollary holds. Consider any $(r, r') \in \Omega_{\mathcal{X},D}$ such that $\mathcal{R}$ is continuous at $r, r'$. Since $(r, r') \in \Omega_{\mathcal{X},D}$, there exist $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ such that $r = |\boldsymbol{u}|$, $r' = |\boldsymbol{u}'|$, $\boldsymbol{u} - \boldsymbol{u}' \in \mathcal{V}_{\mathcal{X}}$ and $|\boldsymbol{u} - \boldsymbol{u}'| \leq D$. By the circularity of $\mathcal{F}_{\mathcal{R}}$, note also that it is continuous at $\boldsymbol{u}, \boldsymbol{u}'$ since its radial $\mathcal{R}$ is continuous at $r, r'$. Now since $\mathcal{F}_{\mathcal{R}}$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$, it holds by Theorem 9 that $\mathcal{F}_{\mathcal{R}}(\boldsymbol{u}) \leq e^{\epsilon} \mathcal{F}_{\mathcal{R}}(\boldsymbol{u}')$, which implies that $\mathcal{R}(r) \leq e^{\epsilon} \mathcal{R}(r')$.

Conversely, assume that the condition of Corollary 12 holds for $\mathcal{F}_{\mathcal{R}}$. This means that for all $(r, r') \in \Omega_{\mathcal{X},D}$ such that $\mathcal{R}$ is continuous at $r, r'$, we have $\mathcal{R}(r) \leq e^{\epsilon}\, \mathcal{R}(r')$. We now demonstrate that $\mathcal{F}_{\mathcal{R}}$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$. Consider any pair of points $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ at which $\mathcal{F}_{\mathcal{R}}$ is continuous, $\boldsymbol{u} - \boldsymbol{u}' \in \mathcal{V}_{\mathcal{X}}$, and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') \leq D$. Let $r = |\boldsymbol{u}|$ and $r' = |\boldsymbol{u}'|$. We have that $(r, r') \in \Omega_{\mathcal{X},D}$ and $\mathcal{R}$ is continuous at $r, r'$ since $\mathcal{F}_{\mathcal{R}}$ is circular and continuous at $\boldsymbol{u}, \boldsymbol{u}'$. Therefore it holds that $\mathcal{R}(r) \leq e^{\epsilon}\, \mathcal{R}(r')$, which implies by Definition 11, that $\mathcal{F}_{\mathcal{R}}(\boldsymbol{u}) \leq e^{\epsilon}\, \mathcal{F}_{\mathcal{R}}(\boldsymbol{u}')$. Thus by Theorem 9, $\mathcal{F}_{\mathcal{R}}$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$. □

The above corollary provides a simple way to prove or disprove the satisfiability of ($D$, $\epsilon$)-location privacy for individual circular noise functions. In particular, if $\mathcal{X} = \mathbb{E}^2$, it is clear that $\Omega_{\mathcal{X},D}$ would consist of every pair of distances $(r, r')$ such that $|r - r'| \leq D$. More generally, if $\mathcal{X}$ is a circular region in $\mathbb{E}^2$ with diameter $W_{\mathcal{X}}$ (*e.g.*, covering a particular city or country), it is easy to see that $\Omega_{\mathcal{X},D}$ would consist of all pairs $(r, r')$ having $|r - r'| \leq D_0$ in which $D_0 = \min(D, W_{\mathcal{X}})$. For this situation, we provide concrete examples of circular noise functions that satisfy ($D$, $\epsilon$)-location privacy assuming $D \leq W_{\mathcal{X}}$.

**Example 13.** (*The Laplacian function.*) Consider the function $\mathcal{R}^L_b(r) = b^2/(2\pi)\, e^{-b\, r}$ with the parameter $b$. It is easy to check that $\mathcal{R}^L_b$ is a valid radial since it satisfies the total probability law (5). Furthermore if $b = \epsilon/D$ then all $(r, r') \in \Omega_{\mathcal{X},D}$ (*i.e.*, $|r - r'| \leq D$) satisfy that $\mathcal{R}^L_b(r) \leq e^{\epsilon} \mathcal{R}^L_b(r')$. Hence, by Corollary 12, the radial $\mathcal{R}^L_b$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$ if $b = \epsilon/D$. The circular noise function corresponding to $\mathcal{R}^L_b$ is exactly the *planar Laplacian* function originally introduced in the setting of geo-indistinguishability [7]. Figure 2(a) shows the plot of $\mathcal{R}^L_b$ that satisfies (200m ,1.0)-location privacy (*i.e.*, $b = 1.0/200$).

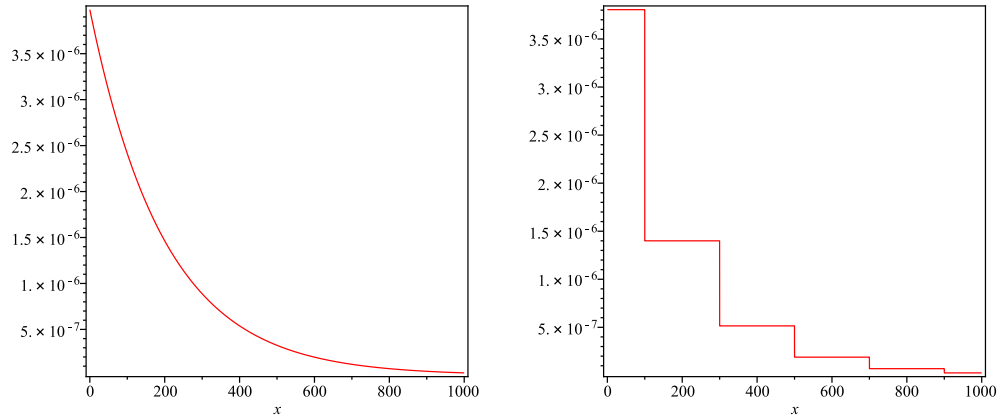
**Example 14.** (*The Stepping function.*) Given the parameters $D > 0$, $s \in [0, D)$, and $\epsilon > 0$, we define the *stepping* noise function $\mathcal{F}_{D,s,\epsilon}$ as the circular noise function having the following radial.

$$\mathcal{R}_{D,s,\epsilon}(r) = \begin{cases} \mathcal{R}_{D,s,\epsilon}(0) & \text{if} \quad 0 \le r < s, \\ e^{-\epsilon}\,\mathcal{R}_{D,s,\epsilon}(0) & \text{if} \quad s \le r < D, \\ e^{-\epsilon}\,\mathcal{R}_{D,s,\epsilon}(r-D) & \text{if} \quad D \le r. \end{cases} \tag{6}$$

To satisfy the total probability law (Equation (5)) the (scaling) constant $\mathcal{R}_{D,s,\epsilon}(0)$ is set as

$$\mathcal{R}_{D,s,\epsilon}(0) = \frac{(1-e^{-\epsilon})^2}{\pi\,(s^2\,(1-e^{-\epsilon})^2 + 2\,s\,e^{-\epsilon}\,D\,(1-e^{-\epsilon}) + e^{-\epsilon}D^2(1+e^{-\epsilon}))}.$$

Note that $\mathcal{R}_{D,s,\epsilon}(r)$ is uniform almost everywhere and discontinuous only at $r = s + k\,D$ for all $k \in \mathbb{N}$. At these values of $r$, the radial $\mathcal{R}_{D,s,\epsilon}(r)$ drops by the factor $e^{-\epsilon}$, taking the form of staircase steps as plotted in Figure 2(b). Therefore, it is straightforward to see that $\mathcal{R}_{D,s,\epsilon}$ satisfies the condition of Corollary 12, and thus satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$. This property holds for $\mathcal{R}_{D,s,\epsilon}$ with any value of $s$ in $[0, D)$, thus enabling the setting of $s$ to be arbitrary. Later in Section 7.2, we demonstrate that by *tuning* this parameter, the stepping noise function provides a better utility than the Laplacian function described previously in Example 13.




(a) Laplacian with $b = 1/200$ (b) Stepping with $D = 200m$, $s = 100m$, $\epsilon = 1.0$

Figure 2: Radials of the Laplacian and Stepping noise functions.

Now that we have described the necessary and sufficient conditions for satisfying ($D$, $\epsilon$)-location privacy for circular noise functions, we can analyze the expected loss of such functions. Let $\mathcal{F}_{\mathcal{R}}$ be a circular noise function whose radial is $\mathcal{R}$. From Proposition 10, we can easily write the expected loss for $\mathcal{F}_{\mathcal{R}}$ as

$$\Psi(\mathcal{F}_{\mathcal{R}}, \mathcal{L}) = \int_0^{\infty} \mathcal{R}(r)\,\mathcal{L}(r)\,2\pi r\,dr. \tag{7}$$

While the expected loss of a circular function does not depend on the domain of interest $\mathcal{X}$, its privacy conditions do. More precisely, these conditions depend on the set $\Omega_{\mathcal{X},D}$ which is determined by the geometry of the domain $\mathcal{X}$. As discussed earlier, when $\mathcal{X}$ is a circular disk with diameter $W_{\mathcal{X}}$, the set $\Omega_{\mathcal{X},D}$ contains exactly every pair $(r, r')$ with $|r - r'| \le D_0$ in which $D_0 = \min(D, W_{\mathcal{X}})$. This highlights an important benefit of using circular noise functions when the domain $\mathcal{X}$ is circular, which is that circular noise functions are rich

enough to cover the full range of expected loss values that are attainable for certain privacy parameters ($D$, $\epsilon$). More precisely, we demonstrate through the following theorem that any noise function satisfying ($D$, $\epsilon$)-location privacy and providing a certain expected loss value can be replaced by a circular one while preserving the same privacy guarantees and the same expected loss value.

**Theorem 15** (Generality of circular noise functions)**.** *Let $\mathcal{X}$ be a circular region. For every noise function $\mathcal{F}$ satisfying ($D$, $\epsilon$)-location privacy for $\mathcal{X}$ and for every loss function $\mathcal{L}$, there exists a circular noise function $\mathcal{F}_\mathcal{R}$ also satisfying ($D$, $\epsilon$)-location privacy and such that $\Psi(\mathcal{F}_\mathcal{R}, \mathcal{L}) = \Psi(\mathcal{F}, \mathcal{L})$.*

*Proof.* Assume that $\mathcal{F}$ satisfies ($D$, $\epsilon$)-location privacy for a circular region $\mathcal{X}$ with diameter $W_\mathcal{X}$. In this case, $\mathcal{V}_\mathcal{X}$ contains every vector having magnitude at most $D_0 = \min(D, W)$ (regardless of its direction). Therefore, it holds by Theorem 9, that all continuity points $\boldsymbol{u}$ and $\boldsymbol{u}' \in \mathbb{E}^2$ with $\mathbf{d}(\boldsymbol{u}, \boldsymbol{u}') \leq D_0$ satisfy $\mathcal{F}(\boldsymbol{u}) \leq e^\epsilon \mathcal{F}(\boldsymbol{u}')$. Also by Definition 1, we can assume without loss of generality that this inequality holds also for discontinuous points.

Using the polar coordinates system, we express each point $\boldsymbol{u} \in \mathbb{E}^2$ by $(r, \theta)$, denoting respectively the distance between $\boldsymbol{u}$ and the origin $\boldsymbol{o}$, and the angular displacement of $\mathbf{vec}(\boldsymbol{u})$ from a reference ray extending from $\boldsymbol{o}$. Then, we define the following functions $\mathcal{F}_\mathcal{R} : \mathbb{E}^2 \to \mathbb{R}^+$ and $\mathcal{R} : [0, \infty) \to \mathbb{R}^+$ such that for all $r \in [0, \infty), \phi \in [0, 2\pi)$

$$\mathcal{F}_\mathcal{R}(r, \phi) = \mathcal{R}(r) = 1/(2\pi) \int_0^{2\pi} \mathcal{F}(r, \theta)\, d\theta. \tag{8}$$

We show in the following that $\mathcal{F}_\mathcal{R}$ is a circular noise function with a radial $\mathcal{R}$. Clearly, $\mathcal{F}_\mathcal{R}(r, \phi)$ depends only on $r$, and is therefore circular. In addition, $\mathcal{F}_\mathcal{R}(r, \phi)$ is the average of $\mathcal{F}$ over all points at distance $r$ from $\boldsymbol{o}$ and therefore $\mathcal{F}_\mathcal{R}$ is bounded since $\mathcal{F}$ is also bounded. Furthermore, $\mathcal{F}_\mathcal{R}$ is a probability density function since $\iint_{\mathbb{E}^2} \mathcal{F}_\mathcal{R}(\boldsymbol{p})\, d\lambda(\boldsymbol{p}) = \int_0^\infty \mathcal{R}(r)\, 2\pi r\, dr = \int_0^\infty \left(1/(2\pi) \int_0^{2\pi} \mathcal{F}(r, \theta)\, d\theta\right) 2\pi r\, dr = 1$, in which the last equality holds because $\mathcal{F}$ is a probability density function.

We now demonstrate that $\mathcal{F}_\mathcal{R}$ is discontinuous only on finitely many analytic curves in any bounded region in $\mathbb{E}^2$. By the circularity of $\mathcal{F}_\mathcal{R}$, its discontinuities occur in circles centered at $\boldsymbol{o}$. Since any bounded region captures a bounded range of distances from $\boldsymbol{o}$, it is enough to prove that only finitely many of these circles exist between any two circles centered at $\boldsymbol{o}$.

For $r \in [0, \infty)$, let $C_r$ denote the circle centered at $\boldsymbol{o}$ and with radius $r$. In addition, we denote by $S_{r_1, r_2}$ the region between $C_{r_1}$ and $C_{r_2}$, for $r_1 < r_2$. By Definition 1, $S_{r_1, r_2}$ contains finitely many analytic curves on which $\mathcal{F}$ is discontinuous. Let $K_r$ be any of these curves that intersects with a circle $C_r \subset S_{r_1, r_2}$ in an arc $A_r$ (of length $\mathfrak{L}(A_r) > 0$). It follows from [18, Corollary 26.5] that $K_r$ must entirely lie on $C_r$. In this situation, we call $C_r$ a discontinuity circle of $\mathcal{F}$. Since the discontinuity curves of $\mathcal{F}$ inside $S_{r_1, r_2}$ are finite, there must be finitely many discontinuity circles of $\mathcal{F}$ in $S_{r_1, r_2}$ for any $r_1 < r_2$.

Now let $\mathcal{D}$ be the set of all discontinuity circles of $\mathcal{F}$ in $\mathbb{E}^2$, and consider any $C_r \notin \mathcal{D}$. We demonstrate that $\mathcal{F}_\mathcal{R}$ is continuous on $C_r$ or equivalently that $\mathcal{R}$ is continuous at $r$. Let $K_1, K_2, \ldots, K_n$ be the discontinuity curves of $\mathcal{F}$ intersecting with $C_r$ respectively at angles $\theta_1, \theta_2, \ldots, \theta_n$. Since $\mathcal{D}$ has only finitely many elements between any two circles, we can find $r_0 \neq r$ such that $C_{r_0} \notin \mathcal{D}$ and no element of $\mathcal{D}$ is between $C_r, C_{r_0}$. We now set $r_0$ to be sufficiently close to $r$ such that every curve $K_i$ intersects with every circle between $C_r$, $C_{r_0}$ inclusive in a single point. Let $\langle C_{r_k} : k \in \mathbb{N} \rangle$ be any sequence of circles between $C_r, C_{r_0}$ converging to $C_r$ as $k \to \infty$. Since $\mathcal{F}$ is continuous on every circle $C_{r_k}$ except the intersection

points with curves $K_i$, $\mathcal{F}$ is measurable on the domain $[0, 2\pi)$ for every $k \in \mathbb{N}$. In addition, $\mathcal{F}$ is continuous at all points of $C_r$ except at the intersection points with $K_1, K_2, \ldots, K_n$. Therefore, the values of $\mathcal{F}$ on $C_{r_k}$ converge pointwise almost everywhere in $[0, 2\pi)$ to its values on $C_r$, which means that $\mathcal{F}(r_k, \theta) \to \mathcal{F}(r, \theta) : \forall \theta \in [0, 2\pi) \setminus \{\theta_1, \theta_2, \ldots, \theta_n\}$. As $\mathcal{F}$ is also bounded, it holds by the dominated convergence theorem that $\lim_{k\to\infty} \mathcal{R}(r_k) = 1/(2\pi) \int_0^{2\pi} \lim_{k\to\infty} \mathcal{F}(r_k, \theta)\, d\theta = 1/(2\pi) \int_0^{2\pi} \mathcal{F}(r, \theta)\, d\theta = \mathcal{R}(r)$. As a consequence, we have that $\mathcal{R}$ is continuous at $r$.

We now demonstrate that $\mathcal{F}_\mathcal{R}$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$. First, observe that $\Omega_{\mathcal{X},D}$ consists exactly of all pairs $(r, r')$ satisfying $|r - r'| \leq D_0$. For any of these pairs, it holds by the properties of $\mathcal{F}$ that $\mathcal{F}(r, \theta) \leq e^\epsilon\, \mathcal{F}(r', \theta)$, $\forall \theta \in [0, 2\pi)$. By integrating this inequality over $[0, 2\pi)$, scaling by $1/2\pi$, and using Equation (8) we get that $\mathcal{R}(r) \leq e^\epsilon\, \mathcal{R}(r')$. Therefore, by Corollary 12, $\mathcal{F}_\mathcal{R}$ satisfies ($D$, $\epsilon$)-location privacy for $\mathcal{X}$. Finally, it holds that $\Psi(\mathcal{F}_\mathcal{R}, \mathcal{L}) = \int_0^\infty \mathcal{R}(r)\, \mathcal{L}(r)\, 2\pi r\, dr = \int_0^\infty \mathcal{L}(r) \int_0^{2\pi} \mathcal{F}(r, \theta)\, r\, d\theta\, dr = \Psi(\mathcal{F}, \mathcal{L})$. □

The above result is fundamental for designing mechanisms that preserve the location privacy with given values of $D$ and $\epsilon$. In particular, this theorem shows that for circular regions $\mathcal{X}$, there is no advantage gained by using a non-circular noise function with respect to privacy and utility. Equivalently, any non-circular noise function for a symmetric mechanism can be substituted by a circular one while preserving the same privacy guarantees and the same expected loss value. This observation simplifies the issue of designing a privacy mechanism to identifying the appropriate circular noise function (or equivalently its radial) using Corollary 12. Hereafter, we describe an algorithm for efficiently implementing a circular noise function.

## 5.2 Implementation of Circular Noise Functions

We consider a symmetric mechanism whose noise function $\mathcal{F}$ is circular. Recall that the output of this mechanism can be viewed as the sum of the real location (*i.e.*, a point in $\mathcal{X}$) and a random noise vector $\vec{\mu}$. Furthermore, $\vec{\mu}$ is the position vector of some point $\boldsymbol{u} \in \mathbb{E}^2$ sampled probabilistically according to the noise function $\mathcal{F}$. Thus, we basically need a procedure to sample $\boldsymbol{u}$.

Relying on the circularity aspect of $\mathcal{F}$, we use the polar coordinates system to describe the points in $\mathbb{E}^2$. Each point $\boldsymbol{u}$ is described by a tuple $(x, \theta)$, in which $x$ is the distance $\mathbf{d}(\boldsymbol{o}, \boldsymbol{u})$ between the origin point $\boldsymbol{o}$ and $\boldsymbol{u}$, and $\theta$ is the angle between the ray extending from $\boldsymbol{o}$ to $\boldsymbol{u}$ and a fixed (reference) ray extending from $\boldsymbol{o}$. We describe these two coordinates by the random variables $X$ and $\Theta$ respectively. Thus, we can express the *joint cumulative distribution function* (joint CDF) of $(X, \Theta)$ as

$$\mathcal{C}_{(X,\Theta)}(x, \theta) = P(X \leq x, \Theta \leq \theta) = \int_0^x \int_0^\theta \mathcal{R}(x) x\, d\theta\, dx.$$

From this expression, one can easily compute the *joint probability density function* (joint pdf) $f_{(X,\Theta)}$ for $(X, \Theta)$ as the partial derivative of the joint CDF $\mathcal{C}_{(X,\Theta)}(x, \theta)$ with respect to $x, \theta$ (in any order):

$$f_{(X,\Theta)}(x, \theta) = \frac{\partial^2 \mathcal{C}_{(X,\Theta)}(x, \theta)}{\partial x\, \partial \theta} = x\, \mathcal{R}(x). \tag{9}$$

We can also compute the "marginal" pdfs for $X$ and $\Theta$ (denoted respectively by $f_X$ and $f_\Theta$) by integrating the joint pdf $f_{(X,\Theta)}(x, \theta)$ respectively on $\theta$ (from 0 to $2\pi$) and $x$ (from 0 to $\infty$). Thus

$$f_X(x) = \mathcal{R}(x)\, 2\pi x, \qquad f_\Theta(\theta) = 1/2\pi. \tag{10}$$

This second equation follows directly from Equation (5). From the above two equations, we can draw the following conclusion.

$$f_{(X,\Theta)}(x,\theta) = f_X(x)\, f_\Theta(\theta),$$

which means that the random variables $X$ and $\Theta$ are independent from each other. This observation implies that they can be sampled independently using their respective pdfs given by Equations (10). Sampling the values of $\Theta$ is easy since its pdf $f_\Theta(\theta)$ is uniform over $[0, 2\pi]$. However, sampling from $X$ is non-trivial and requires the application of the well-known *inverse transform sampling* that relies on the cumulative distribution function $\mathcal{C}_X$ of $X$. This procedure works by drawing a random number $y$ uniformly from $[0, 1]$, and then selecting $x$ such that $\mathcal{C}_X(x) = y$; *i.e.* $x = \mathcal{C}_X^{-1}(y)$. This point-sampling method is generic as it can be applied to arbitrary circular noise functions. For instance, it coincides with the procedure proposed in [7] for drawing points from the planar Laplacian noise function.

Based on the above discussion, we propose Algorithm 1 to implement a symmetric mechanism whose underlying noise function is circular with the radial $\mathcal{R}$.

**Algorithm 1:** Obfuscation Mechanism with a Radial Function $\mathcal{R}$

**Data**: the radial function $\mathcal{R}$ and the real location of the user $\boldsymbol{i}$;
Let $C_X(y) = \int_0^y \mathcal{R}(x)\, 2\pi x\, dx$
**Result**: the obfuscated location $\boldsymbol{p} = \mathcal{K}(\boldsymbol{i})$
1 Draw a number $\theta$ uniformly from $[0, 2\pi]$ ;
2 Draw a number $y$ uniformly from $[0, 1]$ and set $x = C_X^{-1}(y)$ ;
3 Let $\boldsymbol{u}$ be the point with the polar coordinates $(x, \theta)$ ;
4 Let $\vec{\mu}$ be the position vector of $\boldsymbol{u}$ (*i.e.*, $\vec{\mu} = \mathbf{vec}(\boldsymbol{u})$) ;
5 **return** the point $\boldsymbol{p} = \boldsymbol{i} + \vec{\mu}$.

While the output domain of our algorithm (mechanism) is the Euclidean space, which is continuous, in practice the service provider may expect values from a discrete domain $\mathcal{Z}$ (*e.g.*, finite-precision latitude/longitude coordinates, or city names). In this situation, a natural approach is to map the output $\boldsymbol{p}$ of the mechanism to the appropriate element in $\mathcal{Z}$. As we explained in Section 3.5, this post-processing step does not give any new information about the real location of the user as the post-processing procedure can be defined in advance without taking into account the user's real location. Hence, this post-processing does not affect the privacy guarantees of the mechanism (*cf.* Proposition 5).

# 6 Generalizing to Arbitrary Distinguishability Functions

In the previous sections, we have seen that $(D, \epsilon)$-location privacy restricts the distinguishability level between any two points depending on the distance between them. Specifically if this distance is less than $D$ (*i.e.*, points are adjacent to each other) then the two points should be distinguishable up to the level $\epsilon$. In this section, we move from this specific model to a more general privacy notion in which the distinguishability between any two

points is a generic function $\ell$ of these two points. More precisely, we require a mechanism satisfying the following condition for every two points $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$.

$$P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\ell(\boldsymbol{i},\boldsymbol{j})} P(\mathcal{K}(\boldsymbol{j}) \in S) \quad \forall S \subseteq \mathbb{E}^2.$$

We assume that the distinguishability function $\ell$ is symmetric and non-negative. Furthermore, we make the assumption that the distinguishability $\ell$ for two points depends only on the Euclidean distance between them, and therefore may be written as $\ell : [0, \infty) \to [0, \infty)$. Based on these assumptions on $\ell$, we define a generic notion of location privacy that we coin as $\ell$-privacy.

**Definition 16.** ($\ell$-privacy) A mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfies $\ell$-privacy if and only if for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$:

$$P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\ell(\mathbf{d}(\boldsymbol{i},\boldsymbol{j}))} P(\mathcal{K}(\boldsymbol{j}) \in S) \quad \forall S \subseteq \mathbb{E}^2.$$

The above formulation of $\ell$-privacy is similar to the existing notion of $d_{\mathcal{X}}$-privacy [27], which was generically defined for any domain $\mathcal{X}$ of secrets equipped with a metric $d_{\mathcal{X}}$. However, the two notions differ in two fundamental aspects. First, unlike $d_{\mathcal{X}}$, the distinguishability function $\ell$ for two points is not necessarily a metric. Second, we restrict $\ell$ to depend only on the Euclidean distance between the given points. Depending on the form of $\ell$, we can now instantiate different models of location privacy as shown by the following examples.

**$\epsilon$-Geo-indistinguishability.** One instance of the $\ell$-privacy is $\epsilon$-geo-indistinguishability [7] which is obtained by setting the distinguishability function to be the Euclidean distance scaled by a factor $\epsilon$ (*i.e.*, $\ell(d) = \epsilon\, d$). This allows the distinguishability between every two points to change linearly with the distance between them[5].

**($D$, $\epsilon$)-Location Privacy.** For an arbitrary domain of locations $\mathcal{X}$, the constraints of ($D$, $\epsilon$)-location privacy can be easily seen as an instance of $\ell$-privacy with the distinguishability $\ell$ defined as follows

$$\ell(d) = \{\epsilon \text{ if } d \leq D \text{ and } \infty \text{ otherwise } \}.$$

The above function restricts the distinguishability (to $\epsilon$) between only points that are at most $D$ apart. This means that it satisfies a user who can be arbitrarily located at any point in $\mathcal{X}$ but requires always to have a limited distinguishability between his location and *only* points that are in his $D$-proximity. In this situation, the user is not interested in constraining the distinguishability between his (arbitrary) location and the points (of $\mathcal{X}$) that are at distance more than $D$ from him, making the distinguishability function $\ell(d)$ diverge to $\infty$ for $d > D$. This makes the above function a non-metric measure on $\mathcal{X}$ as it violates the triangle inequality (unlike the case of $\epsilon$-geo-indistinguishability).

While the choice of a distinguishability function can be made independently of the domain $\mathcal{X}$, applying the restriction of this function *recursively* on the points of a specific domain may induce a new distinguishability function. For instance if the domain of locations is set to $\mathbb{E}^2$, then applying the constraints of ($D$, $\epsilon$)-location privacy recursively on adjacent points leads to satisfying a staircase distinguishability function.

[5] While the $\epsilon$-geo-indistinguishability is defined for arbitrary output domains, our results concern only the case in which the output domain of a mechanism is $\mathbb{E}^2$.

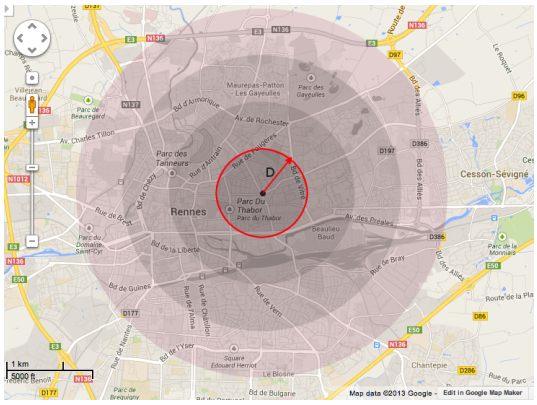



Figure 3: The distinguishability levels provided by $(D, \epsilon)$-location privacy on $\mathbb{E}^2$.

**Proposition 17** ($(D, \epsilon)$-location privacy on $\mathbb{E}^2$)**.** *A mechanism $\mathcal{K} : \mathbb{E}^2 \to \mathbb{E}^2$ satisfies $(D, \epsilon)$-location privacy if and only if for every $\boldsymbol{i}, \boldsymbol{i}' \in \mathbb{E}^2$:*

$$P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\epsilon \lceil \mathbf{d}(\boldsymbol{i},\boldsymbol{i}')/D \rceil} P(\mathcal{K}(\boldsymbol{i}') \in S) \quad \forall S \subseteq \mathbb{E}^2.$$

*Proof.* Let $\mathcal{K}$ be a mechanism satisfying the condition of Proposition 17. For all $\boldsymbol{i}, \boldsymbol{i}'$ : $\mathbf{d}(\boldsymbol{i}, \boldsymbol{i}') \leq D$ and all $S \subseteq \mathbb{E}^2$, the inequality in Definition 2 holds.

Conversely, suppose that $\mathcal{K}$ satisfies the condition of $(D, \epsilon)$-location privacy in Definition 2. Consider any two points $\boldsymbol{i}, \boldsymbol{i}' \in \mathbb{E}^2$. If $\mathbf{d}(\boldsymbol{i}, \boldsymbol{i}') \leq D$ then the inequality of the proposition follows trivially from Definition 2. Otherwise, let $n = \lceil \mathbf{d}(\boldsymbol{i}, \boldsymbol{i}')/D \rceil$. Observe that $n \geq 2$ and $(n-1)D < \mathbf{d}(\boldsymbol{i}, \boldsymbol{i}') \leq nD$. In this case, there are $n-1$ points $\boldsymbol{i}_1, \boldsymbol{i}_2, \ldots, \boldsymbol{i}_{n-1}$ on the line between $\boldsymbol{i}$ and $\boldsymbol{i}'$ such that $\mathbf{d}(\boldsymbol{i}, \boldsymbol{i}_k) = k\,D$ for all $k = 1, 2, \ldots, n-1$. By Definition 2, it holds for any $S \subseteq \mathbb{E}^2$ that $P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\epsilon} P(\mathcal{K}(\boldsymbol{i}_1) \in S) \leq e^{2\epsilon} P(\mathcal{K}(\boldsymbol{i}_2) \in S) \leq \cdots \leq e^{(n-1)\epsilon} P(\mathcal{K}(\boldsymbol{i}_{n-1}) \leq e^{n\epsilon} P(\mathcal{K}(\boldsymbol{i}') \in S)$. Therefore $P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\epsilon\, n} P(\mathcal{K}(\boldsymbol{i}')$, and thus we obtain the condition in the proposition by substituting $n$ using its definition. □

The distinguishability between two points that are $d$-apart follows a staircase function $\epsilon \lceil d/D \rceil$ (in which $\lceil . \rceil$ is the ceiling function). Figure 3 shows the impact of this function on the privacy of the user. The distinguishability level is minimum in the inner-most circular region around the user's location, and increases by steps of length $D$ as we go further away from the user's location. This monotonicity is not unique to the distinguishability function of $(D, \epsilon)$-location privacy but holds for other instances of $\ell$-privacy. For example, the distinguishability function of $\epsilon$-geo-indistinguishability is also increasing with the distance but rather in a linear manner (*i.e.*, $\ell(d) = \epsilon\, d$).

**Extending the formal results to the generic $\ell$-privacy.** Most of our results for $(D, \epsilon)$-location privacy are based on the property that the distinguishability level between two points depends on the distance between them. By relying on this property, which holds for all instances of $\ell$-privacy, we can extend our formal results to the $\ell$-privacy allowing to design mechanisms that satisfy arbitrary distinguishability functions $\ell$ on an arbitrary location domain $\mathcal{X}$.

In this extension, we still model a mechanism by a (probabilistic) mapping from a domain $\mathcal{X} \subseteq \mathbb{E}^2$ to $\mathbb{E}^2$ (as described in Section 3.1), in which the probabilistic behavior of this mapping is precisely described by its randomization functions $\mathcal{F}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \mathcal{X}$. This modeling

generalizes the characterization of ($D$, $\epsilon$)-location privacy (in Theorem 4) to $\ell$-privacy as follows.

**Theorem 18** (Characterization of $\ell$-privacy)**.** *Let $\mathcal{K}$ be a mechanism with randomization functions $\mathcal{F}_{\boldsymbol{i}}$ for all $\boldsymbol{i} \in \mathcal{X}$. Then $\mathcal{K}$ satisfies $\ell$-privacy if and only if for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$, and all $\boldsymbol{p} \in \mathbb{E}^2$ where $\mathcal{F}_{\boldsymbol{i}}, \mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$, it holds that:*

$$\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\ell(\boldsymbol{i},\boldsymbol{j})}\, \mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p}). \tag{11}$$

*Proof.* The proof of this theorem is similar to the proof of Theorem 4 except that we consider every pair of points $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ rather than only the ones satisfying $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$, and we use the distinguishability $\ell(\boldsymbol{i}, \boldsymbol{j})$ instead of $\epsilon$. Consider any $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ and any $\boldsymbol{p} \in \mathbb{E}^2$ such that both the randomization functions $\mathcal{F}_{\boldsymbol{i}}, \mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$. Applying Definition 16 to any ball $B_\delta(\boldsymbol{p})$, with radius $\delta > 0$, yields $\bar{\mathcal{F}}^{\boldsymbol{p}}_{\boldsymbol{i}}(\delta) \leq e^{\ell(\boldsymbol{i},\boldsymbol{j})}\, \bar{\mathcal{F}}^{\boldsymbol{p}}_{\boldsymbol{j}}(\delta)$. Afterwards, taking the limits in this inequality when $\delta \to 0$, and substituting the limits using Equation (2) yield Inequality (11). Conversely, assume that Inequality (11) holds for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ and $\boldsymbol{p} \in \mathbb{E}^2$ in which $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$. Since $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous almost everywhere (according to Definition 1), Inequality (11) holds also almost everywhere in $\mathbb{E}^2$. Thus, it holds for any region $S \subseteq \mathbb{E}^2$ by [21, Proposition 4.3.7 ] that $\iint_S \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p})\, d\lambda(\boldsymbol{p}) \leq \iint_S e^{\ell(\boldsymbol{i},\boldsymbol{j})}\, \mathcal{F}_j(\boldsymbol{p})\, d\lambda(\boldsymbol{p})$, which is the condition of Definition 16. □

The $\ell$-privacy can also be characterized by the gain of an adversary's knowledge through a mechanism in a similar sense to the case of ($D$, $\epsilon$)-location privacy. This gain depends on the distinguishability function $\ell$, and is described by the adversary's prior and posterior probability distributions on the points of interest $\mathcal{I}$ of the user in $\mathcal{X}$.

**Proposition 19** (Impact on the adversary's knowledge)**.** *Let $\mathcal{I}$ be any discrete set of locations. A mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfies $\ell$-privacy if and only if it holds for all $\mathcal{I} \subseteq \mathcal{X}$ and all distributions $\pi_{\mathcal{I}}(.)$ on $\mathcal{I}$ that*

$$\pi_{\mathcal{I}}(\boldsymbol{i} \mid S)/\pi_{\mathcal{I}}(\boldsymbol{j} \mid S) \leq e^{\ell(\boldsymbol{i},\boldsymbol{j})}\pi_{\mathcal{I}}(\boldsymbol{i})/\pi_{\mathcal{I}}(\boldsymbol{j}) \qquad \forall \boldsymbol{i}, \boldsymbol{j} \in \mathcal{I}, \forall S \subseteq \mathbb{E}^2.$$

Note from the above result that observing the output of the mechanism magnifies the ratio between the probabilities of two locations by a factor that is bounded by $e^{\ell(\boldsymbol{i},\boldsymbol{j})}$. This means that the distinguishability function for two points restricts the impact of the public output of the mechanism on the adversary's knowledge.

Furthermore, we also show that any post-processing that is independent of the real location of the user does not endanger the $\ell$-privacy of the user in the same manner as described by Proposition 5.

**Proposition 20** (Post-processing preserves privacy)**.** *Consider a mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ satisfying $\ell$-privacy. Let also $\mathcal{M} : \mathbb{E}^2 \to \mathcal{Z}$ be a probabilistic mapping function where $\mathcal{Z}$ is an arbitrary domain. Then, it holds for all $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ that*

$$P\left((\mathcal{M} \circ \mathcal{K})(\boldsymbol{i}) \in Z\right) \leq e^{\ell(\boldsymbol{i},\boldsymbol{j})}\, P\left((\mathcal{M} \circ \mathcal{K})(\boldsymbol{j}) \in Z\right) \quad \forall Z \subseteq \mathcal{Z}.$$

*Proof.* The proof of this proposition is similar to the proof of Proposition 5 except that we consider every pair of points $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ rather than only ones satisfying $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) \leq D$, and we use the distinguishability $\ell(\boldsymbol{i}, \boldsymbol{j})$ instead of $\epsilon$. □

The above three results do not require the assumption that $\ell$ for two points depends only on the distance between them. However, this assumption is fundamental for the following properties of symmetric mechanisms since in this situation we require a single noise function to guarantee the privacy for all points of $\mathcal{X}$. We begin to describe these properties by the following general characterization.

**Theorem 21** (($\ell$-private symmetric mechanisms)**.** *A symmetric mechanism $\mathcal{K} : \mathcal{X} \to \mathbb{E}^2$ with a noise function $\mathcal{F} : \mathbb{E}^2 \to \mathbb{R}^+$ satisfies $\ell$-privacy if and only if for all points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous, and $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}$, it holds that: $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(\mathbf{d}(\boldsymbol{u},\boldsymbol{v}))}\, \mathcal{F}(\boldsymbol{v})$.*

*Proof.* We prove that for a symmetric mechanism having the noise function $\mathcal{F}$, the condition of $\ell$-privacy in Theorem 18 is equivalent to the condition stated by Theorem 21. More precisely we want to prove that the following statements are equivalent.

1. $\forall \boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \forall \boldsymbol{p} \in \mathbb{E}^2$ such that $\mathcal{F}_{\boldsymbol{i}}, \mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$, it holds that $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\ell(\mathbf{d}(\boldsymbol{i},\boldsymbol{j}))}\, \mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p})$.

2. $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ such that $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}$; $\mathcal{F}$ is continuous at $\boldsymbol{u}, \boldsymbol{v}$, it holds that $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(\mathbf{d}(\boldsymbol{u},\boldsymbol{v}))}\, \mathcal{F}(\boldsymbol{v})$.

We assume that Statement 1 holds and show that Statement 2 follows. Consider any pair of points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous, $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}$. We show that it must hold for them that $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(\mathbf{d}(\boldsymbol{u},\boldsymbol{v}))}\, \mathcal{F}(\boldsymbol{v})$. Let $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \boldsymbol{p} \in \mathbb{E}^2$ be any points satisfying $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$ and $\boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$. These points exist for $\boldsymbol{u}, \boldsymbol{v}$ as follows. Since $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}$, there exist $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}$ in which $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{i} - \boldsymbol{j}$. Now let $\boldsymbol{p} = \mathbf{vec}(\boldsymbol{i}) + \boldsymbol{u} = \mathbf{vec}(\boldsymbol{j}) + \boldsymbol{v}$. It is easy to see that $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{p}$ satisfy that $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i})$ and $\boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$.
Since $\mathcal{F}$ is continuous at $\boldsymbol{u}, \boldsymbol{v}$ it holds by Lemma 8 that $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at the point $\boldsymbol{p}$. Observe also that $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) = |\boldsymbol{i} - \boldsymbol{j}| = |\boldsymbol{v} - \boldsymbol{u}| = \mathbf{d}(\boldsymbol{u}, \boldsymbol{v})$. From Statement 1, it follows that $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\ell(\mathbf{d}(\boldsymbol{i},\boldsymbol{j}))}\, \mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p})$, which implies, by Definition 7, and the fact that $\mathbf{d}(\boldsymbol{i}, \boldsymbol{j}) = \mathbf{d}(\boldsymbol{u}, \boldsymbol{v})$, that $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(\mathbf{d}(\boldsymbol{u},\boldsymbol{v}))}\mathcal{F}(\boldsymbol{v})$. Repeating the same argument for all pairs $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous and $\boldsymbol{v} - \boldsymbol{u} \in \mathcal{V}_{\mathcal{X}}$, we conclude that Statement 2 holds.
Now, we assume that Statement 2 holds and show that 1 is implied. Consider any triplet of points $\boldsymbol{i}, \boldsymbol{j} \in \mathcal{X}, \boldsymbol{p} \in \mathbb{E}^2$ in which $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are both continuous at $\boldsymbol{p}$. We show that it must hold for them that $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\ell(\mathbf{d}(\boldsymbol{i},\boldsymbol{j}))}\mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p})$. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^2$ be any points satisfying $\boldsymbol{u} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{i}), \boldsymbol{v} = \mathbf{pnt}(\boldsymbol{p} - \boldsymbol{j})$. Such points always exist (and are unique) for $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{p}$. Since $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$, it must hold, by Lemma 8, that $\mathcal{F}$ is continuous at $\boldsymbol{u}, \boldsymbol{v}$. Note also that $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{i} - \boldsymbol{j} \in \mathcal{V}_{\mathcal{X}}$ and $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) = |(\boldsymbol{p} - \boldsymbol{i}) - (\boldsymbol{p} - \boldsymbol{j})| = |\boldsymbol{i} - \boldsymbol{j}| = \mathbf{d}(\boldsymbol{i}, \boldsymbol{j})$. Thus, by Statement 2, it holds for $\boldsymbol{u}, \boldsymbol{v}$ that $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(\mathbf{d}(\boldsymbol{u},\boldsymbol{v}))}\mathcal{F}(\boldsymbol{v})$ which implies, by Definition 7 and the fact that $\mathbf{d}(\boldsymbol{u}, \boldsymbol{v}) = \mathbf{d}(\boldsymbol{i}, \boldsymbol{j})$, that $\mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \leq e^{\ell(\mathbf{d}(\boldsymbol{i},\boldsymbol{j}))}\mathcal{F}_{\boldsymbol{j}}(\boldsymbol{p})$. Repeating the same argument for all triplets $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{p} \in \mathbb{E}^2$ where $\mathcal{F}_{\boldsymbol{i}}$ and $\mathcal{F}_{\boldsymbol{j}}$ are continuous at $\boldsymbol{p}$, we get Statement 1. □

As described in Section 5, symmetric mechanisms can be implemented using circular noise functions. Circular noise functions are individually specified by their radials and provide several interesting features with respect to utility, privacy and the easiness of their implementation.

First, the computation of the expected loss of mechanisms based on circular noise functions is simplified to the integral in Equation (7), which depends only on the radial $\mathcal{R}$ and the loss function $\mathcal{L}$. This integral is computed over the domain $[0, \infty)$, and therefore its evaluation is significantly easier compared to the two-dimensional integral in Proposition (10) that is used to compute the expected loss for an arbitrary noise function.

Second, the necessary and sufficient conditions for a circular function $\mathcal{F}_\mathcal{R}$ to satisfy the $\ell$-privacy are reduced to conditions on its radial function $\mathcal{R}$ and the domain $\mathcal{X}$ in a similar sense to Corollary 12. However, a special care needs to be paid here to the dependency on $\mathcal{X}$. In particular, the set $\Omega_{\mathcal{X},D}$ must be extended to capture the distances from the origin to every pair of points $(\boldsymbol{u}, \boldsymbol{u}')$ satisfying only $\boldsymbol{u} - \boldsymbol{u}' \in \mathcal{V}_\mathcal{X}$ (*i.e.* with no regard to the condition $|\boldsymbol{u} - \boldsymbol{u}'| \leq D$). Therefore, instead of relying on $\Omega_{\mathcal{X},D}$, we use the generalized set $\Omega_\mathcal{X}$ defined as

$$\Omega_\mathcal{X} = \{(|\boldsymbol{u}|, |\boldsymbol{u}'|) : \boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2,\, \boldsymbol{u}' - \boldsymbol{u} \in \mathcal{V}_\mathcal{X}\}.$$

At the same time, the distinguishability (provided by $\mathcal{R}$) for two radii $(r, r')$ in $\Omega_\mathcal{X}$ has to fulfill the distinguishability $\ell$ for every two points $\boldsymbol{u}$ and $\boldsymbol{u}'$ with $r = |\boldsymbol{u}|$ and $r' = |\boldsymbol{u}'|$. This requirement is satisfied by using the minimum value of $\ell$ for all such pairs $(\boldsymbol{u}, \boldsymbol{u}')$. For every $(r, r') \in \Omega_\mathcal{X}$, we define the "minimal" distinguishability level $\ell_\mathcal{X}(r, r')$ in the following manner:

$$\ell_\mathcal{X}(r, r') = \min\{\, \ell(|\boldsymbol{u} - \boldsymbol{u}'|) \,:\, \boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2,\, r = |\boldsymbol{u}|, r' = |\boldsymbol{u}'|,\, (\boldsymbol{u} - \boldsymbol{u}') \in \mathcal{V}_\mathcal{X}\}.$$

We remark that $\ell_\mathcal{X}(r, r')$ can be computed as follows. Consider the two circles with radii $r, r'$ and having the same center. Then by considering every two points $\boldsymbol{u}, \boldsymbol{u}'$ lying respectively on the first and second circles such that $\boldsymbol{u} - \boldsymbol{u}' \in \mathcal{V}_\mathcal{X}$, the value of $\ell_\mathcal{X}(r, r')$ is exactly the minimum value of $\ell(|\boldsymbol{u} - \boldsymbol{u}'|)$. In particular, if $\ell(d)$ is monotonically increasing with $d$, it is easy to see that $\ell_\mathcal{X}(r, r') = \ell(|r - r'|)$.

Now, we can phrase the required conditions on a circular noise function $\mathcal{F}_\mathcal{R}$ to satisfy the general $\ell$-privacy for $\mathcal{X}$ in the following manner.

**Corollary 22** ($\ell$-privacy of circular noise functions)**.** *A circular noise function $\mathcal{F}_\mathcal{R}$ having a radial $\mathcal{R}$ satisfies $\ell$-privacy for a domain $\mathcal{X}$ if and only if for all $(r, r') \in \Omega_\mathcal{X}$ in which $\mathcal{R}$ is continuous at $r, r'$ it holds that $\mathcal{R}(r) \leq e^{\ell_\mathcal{X}(r,r')}\, \mathcal{R}(r')$.*

*Proof.* Assume that $\mathcal{F}_\mathcal{R}$ satisfies $\ell$-location privacy for a domain $\mathcal{X}$. In the following we demonstrate that the condition in the corollary holds. Consider any $(r, r') \in \Omega_\mathcal{X}$ such that $\mathcal{R}$ is continuous at $r, r'$. By the circularity of $\mathcal{F}_\mathcal{R}$, it is continuous at all points $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ satisfying $r = |\boldsymbol{u}|$, $r' = |\boldsymbol{u}'|$. By Theorem 21, it must hold that

$$\mathcal{R}(r) \leq e^{\ell(|\boldsymbol{u} - \boldsymbol{u}'|)}\, \mathcal{R}(r') \quad \forall \boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2 : r = |\boldsymbol{u}|, r' = |\boldsymbol{u}'|, (\boldsymbol{u} - \boldsymbol{u}') \in \mathcal{V}_\mathcal{X},$$

which implies by the definition of $\ell_\mathcal{X}(r, r')$ that $\mathcal{R}(r) \leq e^{\ell_\mathcal{X}(r,r')}\, \mathcal{R}(r')$.

Conversely, assume that the condition of Corollary 22 holds for $\mathcal{F}_\mathcal{R}$. This means that for all $(r, r') \in \Omega_\mathcal{X}$ such that $\mathcal{R}$ is continuous at $r, r'$, we have $\mathcal{R}(r) \leq e^{\ell_\mathcal{X}(r,r')}\, \mathcal{R}(r')$. We demonstrate that $\mathcal{F}_\mathcal{R}$ satisfies $\ell$-privacy for $\mathcal{X}$. Consider any $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ at which $\mathcal{F}_\mathcal{R}$ is continuous, $\boldsymbol{u} - \boldsymbol{u}' \in \mathcal{V}_\mathcal{X}$ and then $(|\boldsymbol{u}|, |\boldsymbol{u}'|) \in \Omega_\mathcal{X}$. Furthermore $\mathcal{R}$ is continuous at $|\boldsymbol{u}|, |\boldsymbol{u}'|$ since $\mathcal{F}_\mathcal{R}$ is circular and continuous at $\boldsymbol{u}, \boldsymbol{u}'$. As a consequence, it holds that $\mathcal{R}(|\boldsymbol{u}|) \leq e^{\ell_\mathcal{X}(|\boldsymbol{u}|,|\boldsymbol{u}'|)}\, \mathcal{R}(|\boldsymbol{u}'|)$. From the definition of $\ell_\mathcal{X}$, it also holds for $\boldsymbol{u}, \boldsymbol{u}'$ that $\ell_\mathcal{X}(|\boldsymbol{u}|, |\boldsymbol{u}'|) \leq \ell(|\boldsymbol{u} - \boldsymbol{u}'|)$. Thus by this inequality and Definition 11, we obtain $\mathcal{F}_\mathcal{R}(\boldsymbol{u}) \leq e^{\ell(|\boldsymbol{u} - \boldsymbol{u}'|)}\, \mathcal{F}_\mathcal{R}(\boldsymbol{u}')$. Finally, $\mathcal{F}_\mathcal{R}$ satisfies $\ell$-privacy for $\mathcal{X}$ by Theorem 21. □

Similar to our analysis of $(D, \epsilon)$-location privacy, the set $\Omega_\mathcal{X}$ (and also the function $\ell_\mathcal{X}$) depends on the geometry of the domain $\mathcal{X}$. For instance in the special case in which $\mathcal{X} = \mathbb{E}^2$, it is easy to see that $\Omega_\mathcal{X}$ consists of each pair of distances $(r, r')$ since $\mathcal{V}_\mathcal{X}$ consists of all Euclidean vectors. More generally if $\mathcal{X}$ is a circular region with diameter $W_\mathcal{X}$, the set $\mathcal{V}$ consists of all vectors of magnitude at most $W_\mathcal{X}$, and the set $\Omega_\mathcal{X}$ is composed of all distances $(r, r')$

having $|r - r'| \leq W_{\mathcal{X}}$. These arguments for the case in which $\mathcal{X}$ is a circular domain lead to an important result extending Theorem 15 to the general $\ell$-privacy. In a nutshell, this result states that circular noise functions are rich enough to cover the full range of expected loss values achievable under the privacy constraints resulting from a given distinguishability function. Stated differently, it means that any noise function satisfying $\ell$-privacy can be replaced by a circular noise function without any loss of privacy or utility.

**Theorem 23** (Generality of circular noise functions)**.** *Let $\mathcal{X}$ be a circular region. Then for every noise function $\mathcal{F}$ satisfying $\ell$-privacy for $\mathcal{X}$ and for every loss function $\mathcal{L}$, there exists a circular noise function $\mathcal{F}_{\mathcal{R}}$ also satisfying $\ell$-privacy, and has $\Psi(\mathcal{F}_{\mathcal{R}}, \mathcal{L}) = \Psi(\mathcal{F}, \mathcal{L})$.*

*Proof.* When $\mathcal{X}$ is circular with diameter $W_{\mathcal{X}}$, the condition of a noise function $\mathcal{F}$ to satisfy $\ell$-privacy for $\mathcal{X}$ is that every $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{E}^2$ at which $\mathcal{F}$ is continuous, and having $|\boldsymbol{u} - \boldsymbol{u}'| \leq W_{\mathcal{X}}$ satisfy $\mathcal{F}(\boldsymbol{u}) \leq e^{\ell(|\boldsymbol{u}-\boldsymbol{u}'|)}\mathcal{F}(\boldsymbol{u}')$ (by Theorem 21). We proceed by using similar lines of arguments as the proof of Theorem 15 and define $\mathcal{F}_{\mathcal{R}} : \mathbb{E}^2 \to \mathbb{R}^+$ and $\mathcal{R} : [0, \infty) \to \mathbb{R}^+$ such that for all $r \in [0, \infty), \phi \in [0, 2\pi)$

$$\mathcal{F}_{\mathcal{R}}(r, \phi) = \mathcal{R}(r) = 1/(2\pi) \int_0^{2\pi} \mathcal{F}(r, \theta)\, d\theta. \tag{12}$$

By the same argument as in the proof of Theorem 15, $\mathcal{F}_{\mathcal{R}}$ is a circular noise function with the radial $\mathcal{R}$.

We now demonstrate that $\mathcal{F}_{\mathcal{R}}$ satisfies $\ell$-privacy for $\mathcal{X}$. First, we remark that $\Omega_{\mathcal{X}}$ consists exactly of all pairs $(r, r')$ satisfying $|r - r'| \leq W_{\mathcal{X}}$. For each of these pairs $(r, r')$, it follows from the fact that $\mathcal{F}$ satisfies $\ell$-privacy, and from the definition of $\ell_{\mathcal{X}}(r, r')$ that there is a fixed $\delta \in [0, 2\pi)$ satisfying $\mathcal{F}(r, \theta) \leq e^{\ell_{\mathcal{X}}(r,r')}\, \mathcal{F}(r', \theta + \delta)$ for all $\theta \in [0, 2\pi)$.

By integrating this inequality with respect to $\theta$ on its range $[0, 2\pi)$, scaling by $1/2\pi$, and using Equation (12), we obtain that $\mathcal{R}(r) \leq e^{\ell_{\mathcal{X}}(r,r')}\, \mathcal{R}(r')$. Thus, by Corollary 22, $\mathcal{F}_{\mathcal{R}}$ satisfies $\ell$-privacy for $\mathcal{X}$. Finally it holds that $\Psi(\mathcal{F}_{\mathcal{R}}, \mathcal{L}) = \Psi(\mathcal{F}, \mathcal{L})$ using the same argument as in the proof of Theorem 15. □

Due to Theorem 23, the design of a mechanism for protecting the location privacy of a user in a bounded region (which may contain a high number of points of interest for the user) is a relatively simple process. First, we denote by $\mathcal{X}$ the minimal disk that covers the region considered. Then, it is sufficient (without loss of utility or privacy) to identify the circular function whose radial satisfies $\ell$-privacy using Corollary 22.

Finally, as discussed in Section 5.2, the nature of circular noise functions provides an easy method to draw points from $\mathbb{E}^2$ using Algorithm 1. This makes it relatively easy to implement the mechanisms based on these functions. The key idea here is that the polar coordinates ($r$, $\theta$) of the output point are probabilistically independent, which means that they can be sampled independently of each other.

# 7 Comparison with Other Probabilistic Metrics of Location Privacy

In this section, we discuss the relationship between our framework and other notions of location privacy such as the expected adversary's error as well as geo-indistinguishability.

## 7.1 Relation to the Expected Adversary's Error

As mentioned in Section 2, the authors of [10] have proposed an intuitive probabilistic metric for measuring the location privacy, which has been used in their subsequent work [8, 26]. This metric quantifies the location privacy as the expected adversary's error with respect to his inferences. In the sporadic case, the adversary uses the reported location $\boldsymbol{p}$ together with his knowledge about both the user's mobility profile $\pi$ and the mechanism $\mathcal{K}$ to make an estimate $\hat{\boldsymbol{p}}$ for the real location of the user. The adversary's error is then measured using a distortion distance $d_p$ between his guess $\hat{\boldsymbol{p}}$ and the real user's location $\boldsymbol{i}$. Denoting the points of interest for the user by $\mathcal{I}$ and using the generic probabilistic function $\mathcal{M} : \mathbb{E}^2 \to \mathcal{I}$ to model the adversary's mapping of the reported location to a member of the set $\mathcal{I}$, this privacy metric can be written in our notations as follows:

$$LP(\pi, \mathcal{K}, \mathcal{M}, d_p) = \sum_{\boldsymbol{i} \in \mathcal{I}} \pi(\boldsymbol{i}) \left( \iint_{\mathbb{E}^2} \mathcal{F}_{\boldsymbol{i}}(\boldsymbol{p}) \sum_{\hat{\boldsymbol{p}} \in \mathcal{I}} P(\mathcal{M}(\boldsymbol{p}) = \hat{\boldsymbol{p}})\, d_p(\boldsymbol{i}, \hat{\boldsymbol{p}})\, d\lambda(\boldsymbol{p}) \right).$$

The distance $d_p$ depends on the objective of the inference attack conducted by the adversary. For example, it may be defined to be the Euclidean distance $\mathbf{d}(\boldsymbol{i}, \hat{\boldsymbol{p}})$ if the adversary wants to approximate the user's real location. It might also be defined to be $\{0$ if $\boldsymbol{i} = \hat{\boldsymbol{p}}$ and $1$ otherwise$\}$ if he wants to identify the real location.

It is clear that the privacy metric LP depends strongly on the adversary's knowledge. In one of the extreme cases, the adversary may obtain some public information allowing him to locate the user. In this case, the LP yields the lowest level of privacy even if the mechanism does not leak any information by itself. This highlights the main difference between this previous work and our framework, in which we quantify the privacy guarantees that are provided by the mechanism *itself* and abstract away from the adversary's knowledge (which is hard to model realistically).

Nonetheless, the metric LP is related to our framework through the utility measure. More precisely, while the LP for a mechanism can go arbitrarily low depending on the adversary's knowledge, it is upper-bounded by the expected loss of the mechanism under certain assumptions. In particular, the form of LP is similar to that of the expected loss $\Psi$ (in Definition 6), with the exception that LP uses the privacy distance $d_p$ instead of the loss function $\mathcal{L}$ and takes into account the adversary's mapping $\mathcal{M}$. If the adversary is *Bayesian* (*i.e.*, he carefully chooses the mapping function $\mathcal{M}$ to minimize the value of LP as demonstrated in [26, Sec. 4.1]), it is straightforward to see that the privacy of the user as measured by LP is upper-bounded by the expected loss $\Psi$ of the mechanism if the loss function $\mathcal{L}$ in the computation of $\Psi$ is taken to be $d_p$. Recall that while the loss function $\Psi$ for arbitrary mechanisms depends on the user profile $\pi$, this dependency vanishes in the case of symmetric mechanisms (*cf.* Proposition 10).

## 7.2 Comparison with $\epsilon$-Geo-indistinguishability

**A comparison between $\ell$-privacy and $\epsilon$-geo-indistinguishability.** By comparing our framework with the recent model of $\epsilon$-geo-indistinguishability [7, 28], we observe that they differ on several aspects despite both being based on the idea of differential privacy. First, our framework is more general in the sense that it captures various forms of constraints on the distinguishability between locations including not only $\epsilon$-geo-indistinguishability but also other variants (*e.g.*, ($D$, $\epsilon$)-location privacy). In addition, we recall that in [7, 28], the possible outputs of a mechanism are assumed to be finite and discrete, thus allowing

the use of linear programming techniques to find the optimal mechanism satisfying $\epsilon$-geo-indistinguishability for a given user with a specific prior [28]. In contrast in our framework, we assume that the output domain of a mechanism is the continuous infinite Euclidean space $\mathbb{E}^2$, which is more generic than discrete domains and leads to several interesting features such as the notion of "symmetric mechanism". Such a mechanism computes its output by adding to the user's real location a random "noise vector" drawn from a specific "noise function". This property reduces the design problem of the mechanism to specifying its underlying noise function and makes the utility of the mechanism independent of the prior knowledge of the user. Furthermore, we proved that under certain assumptions a special type of noise functions, called "circular" are generic enough to provide the same privacy and utility levels provided by other non-circular ones. At the same time, the circular functions allow sampling noise vectors using a simple process.

**A comparison between ($D$, $\epsilon$)-location privacy and $\epsilon$-geo-indistinguishability.** We have shown previously that $\epsilon$-geo-indistinguishability and ($D$, $\epsilon$)-location privacy are instances of the general $\ell$-privacy and that both of them allow a level of distinguishability $\ell(d)$ (between the points of $\mathcal{X}$) that increases with the distance. However, they correspond to different privacy requirements that are implemented by their respective distinguishability functions. On one hand, $\epsilon$-geo-indistinguishability satisfies a user who requires to restrict the distinguishability level between his location and every point in the domain, such that this level increases linearly with the distance. On the other hand, the ($D$, $\epsilon$)-location privacy corresponds to a *relaxed* requirement in which the user wants only to restrict this distinguishability within the surrounding neighborhood (points within distance at most $D$ from him) while at the same time the risk of this distinguishability is uniform from his perspective. Thus, ($D$, $\epsilon$)-location enforces less constraints and is implied by $\epsilon/D$-geo-indistinguishability as follows.

**Proposition 24** (Relation between ($D$, $\epsilon$)-location privacy and $\epsilon$-geo-indistinguishability)**.** *A mechanism $\mathcal{K}$ satisfies ($D$, $\epsilon$)-location privacy if it satisfies $\epsilon/D$-geo-indistinguishability.*

*Proof.* If $\mathcal{K}$ satisfies $\epsilon/D$-geo-indistinguishability then by definition, it holds for all $\boldsymbol{i}, \boldsymbol{i}' \in \mathbb{E}^2$ that $P(\mathcal{K}(\boldsymbol{i}) \in S) \leq e^{\epsilon\, \mathbf{d}(\boldsymbol{i},\boldsymbol{i}')/D}\, P(\mathcal{K}(\boldsymbol{i}') \in S)$. If we choose $\mathbf{d}(\boldsymbol{i}, \boldsymbol{i}') \leq D$ then $\epsilon\, \mathbf{d}(\boldsymbol{i}, \boldsymbol{i}')/D \leq \epsilon$. Thus $\mathcal{K}$ satisfies ($D$, $\epsilon$)-location privacy according to Definition 2. □

We illustrate this proposition using the following example in which we assume that the domain of locations is $\mathbb{E}^2$.

**Example 25** (Illustrative example)**.** Consider a user such as Alice who regularly travels to various places in the world, and therefore needs a privacy guarantee that holds everywhere in $\mathbb{E}^2$. Precisely, for every point $\boldsymbol{i}$, she wants to protect whether she is located at $\boldsymbol{i}$ or at another (nearby) point that is less than 200 meters away from $\boldsymbol{i}$. In particular from her point of view, distinguishing between any two points that are $\leq$ 200m apart would incur the same risk for Alice, while this distinguishability is safe for points that are further. Thus, she sets a uniform distinguishability bound of $1.0$ for points that are less than 200m from each other, and allow points that are further apart to be freely distinguishable. Equivalently, this requires ($D$, $\epsilon$)-location privacy on $\mathbb{E}^2$ with $D = 200m$ and $\epsilon = 1.0$. By Proposition 17, these parameters set the distinguishability level (as a function of the distance $d$) to $\ell(d) = \epsilon\, \lceil d/D \rceil$. According to Proposition 24, Alice's requirement is also satisfied by $\epsilon/D$-geo-indistinguishability, which not only restricts the distinguishability between nearby points (when $d \leq D$) as needed, but also enforces this distinguishability to vary with the distance

according to the function $\ell'(d) = (\epsilon/D)\, d$ making closer points less distinguishable. The plots of $\ell(d)$ and $\ell'(d)$ in Figure 4 show that $\epsilon/D$-geo-indistinguishability enforces stronger constraints than those of $(D, \epsilon)$-location privacy required by Alice in this example.

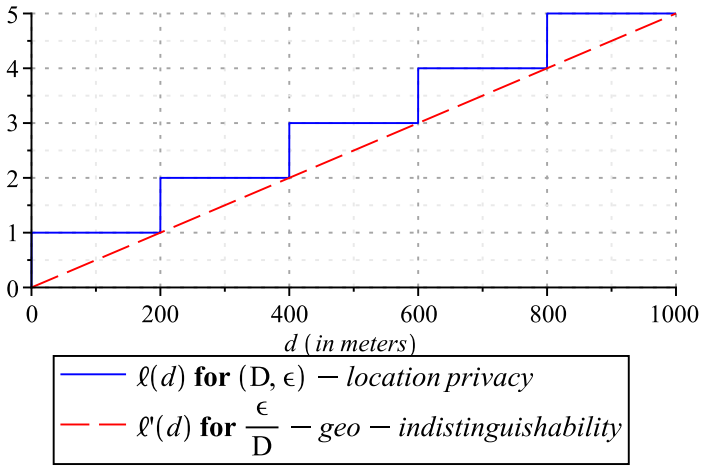


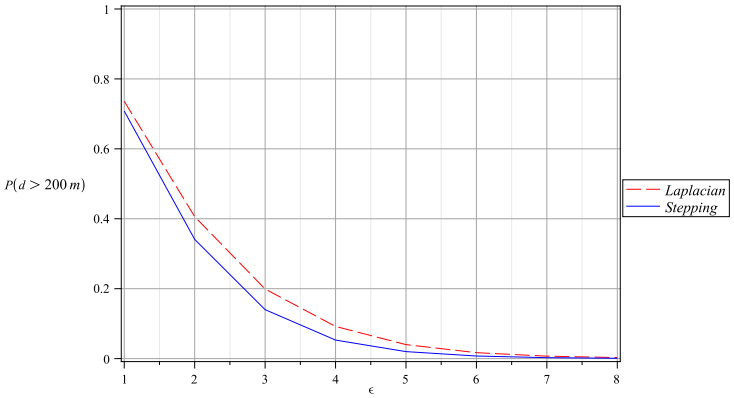

Figure 4: The distinguishability functions $\ell(d)$ and $\ell'(d)$ for example 25.

To summarize, a user asking for $(D, \epsilon)$-location privacy can be simply satisfied by a mechanism ensuring the stronger $\epsilon/D$-geo-indistinguishability. As shown in [7], such a mechanism is typically based on the Laplacian noise function described in Example 13 (*cf.* Section 5). However, since $(D, \epsilon)$-location privacy is more relaxed than $\epsilon/D$-geo-indistinguishability, it is possible to take advantage of this relaxation to gain better utility (*i.e.*, less expected loss). In particular, we observe that for many loss functions, the stepping noise function (described in Example 14) provides always a smaller expected loss compared to the Laplacian one. The rest of this section is dedicated to the demonstration of this observation for a couple of intuitive loss functions: namely, the $\alpha$-*binary loss* and the *distance loss*. In the following, we denote by $d$ the distance between the real location of the user and the reported one.

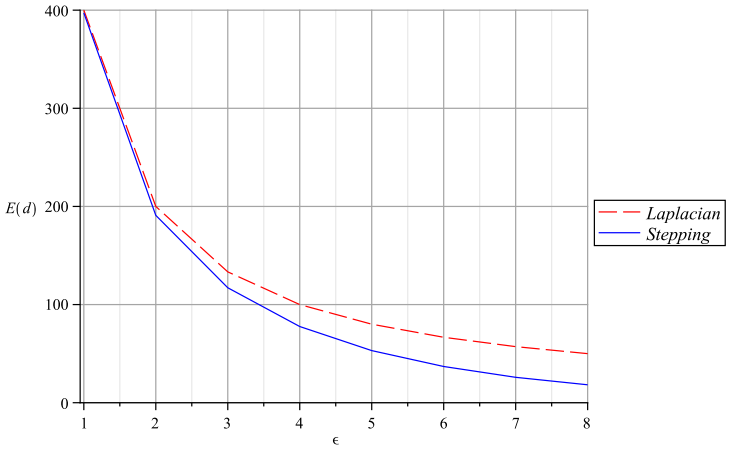
$\alpha$-*binary loss.* This function, denoted by $\mathcal{L}_\alpha^{bin}$, quantifies the loss as 0 if the output of the mechanism is within distance $\alpha$ from the real location, and as 1 otherwise (*i.e.*, $\mathcal{L}_\alpha^{bin}(d) = 0$ if $d \leq \alpha$ and 1 otherwise). Therefore, the expected loss (7) using $\mathcal{L}_\alpha^{bin}$ is exactly the probability $P(d > \alpha)$, which corresponds to the probability of the mechanism to report a point that is at distance $d$, greater than $\alpha$, from the real location. Note that $P(d > \alpha)$ corresponds to the notion of $(\alpha, \delta)$-*usefulness* originally introduced in [29] and also used in [7] to quantify the utility of mechanisms. More precisely, a mechanism $\mathcal{K}$ is said to be $(\alpha, \delta)$-*useful* if $P(d > \alpha) = 1 - \delta$.

Setting the binary loss function to $\alpha = 200m$, we performed a set of experiments to compare between the stepping and the Laplacian noise functions satisfying $(D, \epsilon)$-location privacy for fixed $D = 200m$ and various values of $\epsilon$. For each value of $\epsilon$, the radial $\mathcal{R}_b^L$ of the Laplacian noise function is determined by setting its parameter $b$ to $\epsilon/200$ (as shown in Example 13) and then evaluating the expected loss by Equation (7). The radial of the stepping function has the parameter $s \in [0, D)$ in addition to $D, \epsilon$. We chose the *optimal* value for $s$ by minimizing the expected loss (7). While in general the optimal value of $s$ depends

on $\epsilon$, $D$ and the loss function, we observed that when $D = 200m$ and the loss function is $\mathcal{L}^{bin}_{200}$ (*i.e.*, $\alpha = D$), the optimal value of $s$ is also $200m$ (*i.e.*, $s = \alpha$) for all values of $\epsilon$. This is because when $\alpha \leq D$, the stepping radial (in Figure 2(b)) assigns the highest probability for the event $d \leq \alpha$ if $s = \alpha$, making the probability of the complementary event $d > \alpha$ (*i.e.*, the expected binary loss) minimal. Figure 5(a) shows the results obtained for $\epsilon = 1, 2, \ldots, 8$. From these results, it is clear that the stepping function (*i.e.*, the solid curve) incurs less expected loss compared to the Laplacian one (*i.e.*, the dashed curve). This superiority of the stepping function was observed for all the experiments that we have run for other values of $\alpha$ (*e.g.*, less and greater than $D$). Remark that when $\alpha > D$, the optimal values of $s$ depend on $\epsilon$ and $\alpha$.




(a) Binary loss with $\alpha = 200m$




(b) Distance loss

Figure 5: The expected loss of Laplacian and Stepping noise function with $D = 200m$.

*Distance loss.* This function, denoted by $\mathcal{L}^{dis}$, defines the loss in a single run of the mechanism as the distance $d$ between the real and reported locations (*i.e.*, $\mathcal{L}^{dis}(d) = d$). In this situation, the expected loss is therefore the expected (average) value of such distance. Using the distance loss function, we compare again between the expected losses for the Laplacian and the stepping noise functions satisfying ($200m$, $\epsilon$)-location privacy for various $\epsilon$. For this comparison, we use the same approach as for the binary loss. In particular, while the

Laplacian radial $\mathcal{R}_b^L$ is determined only by the given privacy parameters $\epsilon$ and $D = 200m$ (by setting $b = \epsilon/D$), the stepping function is also determined by the parameter $s$, which we set to its optimal value (for given $D, \epsilon$). Unlike the $\alpha$-binary loss with $\alpha \leq D$, we observed that with the distance loss, the optimal value of $s$ varies according to $\epsilon$. In particular, for $\epsilon = 1, 2, \cdots, 8$ and $D = 200m$, the optimal values for $s$ are found respectively to be $133, 107, 83, 62, 46, 33, 24$, and $17$. Using these values (for $\epsilon$ and the corresponding optimal $s$), we plot in Figure 5(b) the expected distance loss for the two noise functions. For this case also, we observe that the stepping noise (represented by the solid curve) achieves less expected loss than the Laplacian one (*i.e.*, the dashed curve).

Despite the fact that we ran our experiments for $D = 200m$, it can be seen that scaling $D$ by a factor $k > 0$ yields similar results when all distances are also scaled by $k$. Thus, we would obtain the same plot in Figure 5(a) but for the probability $P(d > (200\,k)m)$ instead of $P(d > 200m)$, and also obtain the same curves in Figure 5(b) but with the values of the expected distance loss $E(d)$ scaled by $k$. Consequently, we conclude that for (at least) these two natural loss functions, ($D$, $\epsilon$)-location privacy leads to a smaller expected loss (*i.e.*, better utility) compared to the Laplacian mechanism, which is typically used for ensuring $\epsilon/D$-geo-indistinguishability. The stepping function achieves this improvement provided that its parameter $s$ is optimized for the adopted loss function and the given $D$, $\epsilon$. This optimization of $s$ can easily be performed using standard numerical methods since its value lies in the bounded interval $[0, D)$.

Using the loss functions mentioned previously, one can choose the values of $\epsilon$ depending on the LBS. In particular, many LBSs require the reported location to be relatively precise, for instance by allowing the distance $d$ between the real and reported locations to exceed $D$ with only a negligible probability. Examples for these LBSs include traffic monitoring systems and GPS navigation, in which vehicles submit their positions in a private manner to the server. By looking at Figure 5(a), we can see that for ensuring that $P(d > D) < 0.1$, the value of $\epsilon$ has to be at least $4.0$. From Figure 5(b), we observe that for $\epsilon \geq 4$, adopting ($D$, $\epsilon$)-location privacy using the stepping function provides at least $25\%$ reduction in the expected distance loss compared to the $\epsilon/D$-geo-indistinguishability with the Laplacian one. Thus for these LBSs, we suggest to adopt the former choice with $\epsilon \geq 4$ as a better trade-off between privacy and utility.

However, other LBSs can achieve an acceptable quality with an imprecise reported location, which is further away from the real position. Examples of such LBSs include weather forecasting as well as the retrieval of points of interests. In particular, this last application is handled in [7] by consuming sufficient bandwidth to retrieve all points of interests within a (controllably) large area around the (imprecise) reported point and then locally extract on the user's device the relevant points of interests based on his true location. This type of LBSs allows for higher levels of privacy (*i.e.*, smaller $\epsilon$). For instance with $D = 200m$, it turns out that $P(d > 3D) < 0.1$ is satisfied with $\epsilon \geq 1.3$ for both stepping and Laplacian functions. From Figure 5(b), we also observe that for low values of $\epsilon$ (*i.e.*, for $0 < \epsilon \leq 2$), the stepping function achieves slightly lower expected loss than the Laplacian one.

# 8 Conclusion

Differential privacy [6, 30] was introduced in the context of statistical databases to provide a privacy guarantee for their participants by ensuring that two "adjacent" databases are indistinguishable from each other. In our setting, we ask for a similar requirement in the context of LBSs, which is that two locations are indistinguishable provided that they are

"close" from each other. Based on this analogy, in this work we have adapted differential privacy to the context of location data. This adaptation can lead to various models depending on the precise form of indistinguishability required.

More precisely, we have first introduced the notion of $(D, \epsilon)$-location privacy. Similarly to differential privacy, $(D, \epsilon)$-location privacy has the merit of describing the privacy guarantees in terms of the mechanism itself and abstracts away from the background knowledge that the adversary might have gathered. We have described a model for obfuscation mechanisms working on an arbitrary domain $\mathcal{X}$ of locations, and then used it to characterize $(D, \epsilon)$-location privacy. This characterization is used as the main tool for conducting a subsequent analysis of the mechanisms with respect to privacy and utility. We have also studied a specific class of mechanisms called "symmetric", which can be easily implemented as the output of a symmetric mechanism is produced by adding a random noise vector to the real location. We gave the necessary and sufficient conditions for a noise function to guarantee $(D, \epsilon)$-location privacy for this type of mechanism before deriving an expression for the expected loss of such a mechanism, which has the main advantage of being independent of the prior knowledge.

In addition, we considered the case when the location domain $\mathcal{X}$ is circular and proved in Theorem 15 that circular noise functions (*cf.* Section 5) are sufficient to guarantee the same privacy and utility levels as other (non-circular) noise functions. This result can be used to "squeeze" the design space of noise functions to the subspace of circular ones. We have also extended our results to a generalized notion of location privacy, called $\ell$-privacy capturing both $(D, \epsilon)$-location privacy and also the notion of $\epsilon$-geo-indistinguishability recently introduced by Andrès, Bordenabe, Chatzikokolakis and Palamidessi. Finally, we have compared $\epsilon$-geo-indistinguishability to the more generic notion of $\ell$-privacy as well as to $(D, \epsilon)$-location privacy, in particular with respect to utility and privacy.

In the future, we want to tackle the identification of the optimal (circular) noise function minimizing the expected loss (for a given loss function) while satisfying a desired level of location privacy. Since we believe that Theorem 23 is a useful tool in finding such an optimal mechanism for a region $\mathcal{X}$, we will investigate if this theorem (or a similar one) can capture the cases in which $\mathcal{X}$ is not necessarily circular.

Similarly to differential privacy, if the user applies several times the obfuscation mechanism providing $\ell$-privacy (which is expected to be the case for most users), then he incurs a loss in terms of his privacy budget of $\epsilon$ in the worst case for each of this application, thus slowly depleting his budget. This might not be a problem if the mechanism is used on a sporadic basis (which is the setting that we assumed in this paper) but it might be a critical issue in other situations in which the user has to release his location on a regular basis. To mitigate this issue, it might be possible to use the obfuscation mechanism in combination with other strategies such as an instance the *predictive mechanism* proposed recently by Chatzikokolakis, Palamidessi and Stronati [9]. In a nutshell, this mechanism exploits the correlations in the locations previously revealed in order to guess the new location based on the previously reported locations. If this prediction is successful, then it is unnecessary to call the obfuscation mechanism and it becomes possible to save on the privacy budget (still a small price has to be paid due to the fact that the testing deciding whether the prediction is "good" or not has to be made differentially private). Another possibility when the user often enters in the same area, such as the neighborhood in which he lives, is to report a fixed area instead of using the mechanism. However, one should be careful of how the guarantees provided by the mechanism are affected by the composition with other strategies that do not offer $\ell$-privacy. We are planning to investigate these lines of research as future work.

## Acknowledgements

This work is partially funded by the Inria project lab CAPPRIS (Collaborative Actions on the Protection of Privacy Rights in the Information Society). We would like to thank Catuscia Palamidessi, Kostas Chatzikokolakis and the anonymous reviewers for their comments and recommendations that have greatly help us to improve the paper.

## References

[1] Pfitzmann, A., Köhntopp, M. (2001): Anonymity, unobservability, and pseudonymity - a proposal for terminology. In: *Proceedings of the Workshop on Design Issues in Anonymity and Unobservability*, 1–9.

[2] Gambs, S., Killijian, M., del Prado Cortez, M.N. (2014): De-anonymization attack on geolocated data. *Journal of Computer and Systems Sciences* **80**(8) 1597–1614.

[3] Gruteser, M., Grunwald, D. (2003): Anonymous usage of location-based services through spatial and temporal cloaking. In: *Proceedings of the 1st international conference on Mobile Systems, applications and services (MobiSys'03)*, 31–42.

[4] Gedik, B., Liu, L. (2005): Location privacy in mobile systems: A personalized anonymization model. In: *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, 620–629.

[5] Damiani, M.-L. (2014): Location privacy models in mobile applications: conceptual view and research directions. *GeoInformatica* **18**(4) 819–842.

[6] Dwork, C. (2006): Differential privacy. In: *Proceedings of the 33rd International Conference on Automata, Languages and Programming (ICALP'06)*, 1–12.

[7] Andrés, M.E., Bordenabe, N.E., Chatzikokolakis, K., Palamidessi, C. (2013): Geo-indistinguishability: Differential privacy for location-based systems. In: *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS'13)*, 901–914.

[8] Shokri, R., Theodorakopoulos, G., Danezis, G., Hubaux, J.P., Le Boudec, J.Y. (2011): Quantifying location privacy: The case of sporadic location exposure. In: *Proceedings of the 11th International Conference on Privacy Enhancing Technologies (PETS'11)*, 57–76.

[9] Chatzikokolakis, K., Palamidessi, C., Stronati, M. (2014): A predictive differentially-private mechanism for mobility traces. In: *Proceedings of the 14th International Conference on Privacy Enhancing Technologies (PETS'14)*, 21–41.

[10] Shokri, R., Theodorakopoulos, G., Le Boudec, J.Y., Hubaux, J.P. (2011): Quantifying location privacy. In: *Proceedings of the 2011 IEEE Symposium on Security and Privacy (SP'11)*, 247–262.

[11] Hoh, B., Gruteser, M., Xiong, H., Alrabady, A. (2006): Enhancing security and privacy in traffic-monitoring systems. *IEEE Pervasive Computing* **5**(4) 38–46.

[12] Beresford, A.R., Stajano, F. (2003): Location privacy in pervasive computing. *IEEE Pervasive Computing* **2**(1) 46–55.

[13] Sweeney, L. (2002): $k$-anonymity: A model for protecting privacy. *International Journal Uncertainty, Fuzziness and Knowledge-Based Systems* **10**(5) 557–570.

[14] Mokbel, M.F., Chow, C.Y., Aref, W.G. (2006): The new Casper: qQery processing for location services without compromising privacy. In: *Proceedings of the 32nd international conference on Very large data bases (VLDB'06)*, 763–774.

[15] Shokri, R., Troncoso, C., Diaz, C., Freudiger, J., Hubaux, J.P. (2010): Unraveling an old cloak: $k$-anonymity for location privacy. In: *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society (WPES'10)*, 115–118.

[16] Ho, S.S., Ruan, S. (2011): Differential privacy for location pattern mining. In: *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS (SPRINGL'11)*, 17–24.

[17] Machanavajjhala, A., Kifer, D., Abowd, J., Gehrke, J., Vilhuber, L. (2008): Privacy: Theory meets practice on the map. In: *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering (ICDE'08)*, 277–286.

[18] Agarwal, R., Perera, K., Pinelas, S. (2011): *An Introduction to Complex Analysis*. Springer US.

[19] Halmos, P.R. (1974): *Measure theory*. Springer-Verlag New York.

[20] Aliprantis, C., Burkinshaw, O. (1981): *Principles of real analysis*. North Holland.

[21] Royden, H.L. (1988): *Real analysis* (3rd edition). Macmillan New York.

[22] Gupte, M., Sundararajan, M. (2010): Universally optimal privacy mechanisms for minimax agents. In: *Proceedings of the 29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS'10)*, 135–146.

[23] Brenner, H., Nissim, K. (2010): Impossibility of differentially private universally optimal mechanisms. In: *Proceedings of 51th Annual IEEE Symposium on Foundations of Computer Science (FOCS'10)*, 71–80.

[24] Ghosh, A., Roughgarden, T., Sundararajan, M. (2009): Universally utility-maximizing privacy mechanisms. In: *Proceedings of 41st Annual ACM Symposium on Theory of Computing (STOC'09)*, 351–360.

[25] ElSalamouny, E., Chatzikokolakis, K., Palamidessi, C. (2013): A differentially private mechanism of optimal utility for a region of priors. In: *Proceedings of the Second international conference on Principles of Security and Trust (POST'13)*, 41–62.

[26] Shokri, R., Theodorakopoulos, G., Troncoso, C., Hubaux, J.P., Le Boudec, J.Y. (2012): Protecting location privacy: Optimal strategy against localization attacks. In: *Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS'12)*, 617–627.

[27] Chatzikokolakis, K., Andrés, M.E., Bordenabe, N.E., Palamidessi, C. (2013): Broadening the scope of differential privacy using metrics. In: *Proceedings of the 13th International Conference on Privacy Enhancing Technologies (PETS'13)*, 82–102.

[28] Bordenabe, N.E., Chatzikokolakis, K., Palamidessi, C. (2014): Optimal geo-indistinguishable mechanisms for location privacy. In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS'14)*, 251–262.

[29] Blum, A., Ligett, K., Roth, A. (2008): A learning theory approach to non-interactive database privacy. In: *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC'08)*, 609–618.

[30] Dwork, C. (2011): A firm foundation for private data analysis. *Communications of the ACM* **54**(1) 86–96.